

	1441	TTTCTCTCAACGGTTCTGTCATTCAGGACCAGGATTCA	1480
5		.	.
	1481	CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA	1520
10		.	.
	1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
		.	.
	1561	CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG	1600
15		.	.
	1601	CTTCTGTGACCCCTATTCACCTAACGTTAACGGGTAA	1640
		.	.
	1641	TTCATCCATCTTCTCCAATACAGTTCCAGCTACAGCTACC	1680
20		.	.
	1681	TCCTTGGATAATCTCCAATCCAGCGATTCGGTTACTTG	1720
		.	.
	1721	AAAGTGCCAATGCTTTACATCTTCACTCGGTAACATCGT	1760
25		.	.
	1761	GGGTGTTAGAAACTTTAGTGGGACTGCAGGAGTGATTATC	1800
		.	.
30		GACAGATTGAGTTCAATTCCAGTTACTGCAACACTCGAGG	1840
		.	.
	1801	CTGAGTACAACCTTGAGAGAGGCCAGAAGGCTGTGAACGC	1880
35		.	.
	1881	CCTCTTACCTCCACCAATCAGCTGGTTGAAAACAAAC	1920
		.	.
	1921	GTTACTGACTATCACATTGACCAAGTGTCCAACCTGGTCA	1960
40		.	.
	1961	CCTACCTTAGCGATGAGTTCTGCCTCGACGAGAACCGTGA	2000
		.	.
	2001	ACTCTCCGAGAAAGTTAACACGCCAACGCGTCTCAGCGAC	2040
45		.	.
	2041	GAGAGGAATCTCTTGCAAGACTCCAACCTCAAAGACATCA	2080
		.	.
50		ACAGGCAGCCAGAACGTGGTTGGGTGGAAGCACCGGGAT	2120
		.	.
	2121	CACCATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTAC	2160
		.	.
55		GTCACCCCTCTCCGGAACTTTGACGAGTGCTACCCCTACCT	2200

5	2201	ACTTGTACCAGAAGATCGATGAGTCCAAACTCAAAGCCTT	2240
10	2241	CACCAAGGTATCAACTTAGAGGCTACATCGAACAGACCAA	2280
15	2281	GACCTTGAAATCTACTCGATCAGGTACAATGCCAAGCAG	2320
20	2321	AGACCGTGAATGTCCCAGGTACTGGTCCCTCTGGCCACT	2360
25	2361	TTCTGCCAATCTCCCATTGGAAAGTGTGGAGAGCCTAAC	2400
30	2401	AGATGCGCTCCACACCTTGAGTGGAAATCCTGACTTGGACT	2440
35	2441	GCTCCTGCAGGGATGGCGAGAAGTGTGCCACCATTCTCA	2480
40	2481	TCACTTCTCCTTGGACATCGATGTGGATGTACTGACCTG	2520
45	2521	AATGAGGACCTCGGAGTCTGGGTACATTTCAAGATCAAGA	2560
50	2561	CCCAAGACGGACACGCAAGACTTGGCAACCTTGAGTTCT	2600
55	2601	CGAAGAGAAACCATTGGTCGGTGAAGCTCTCGCTCGTGTG	2640
	2641	AAGAGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAAC	2680
	2681	TCGAATGGGAAACTAACATCGTTACAAGGAGGCCAAGA	2720
	2721	GTCCGTGGATGCTTGTTCGTGAACCTCCAAATATGATCAG	2760
	2761	TTGCAAGCCGACACCAACATGCCATGATCCACGCCAG	2800
	2801	ACAAACGTGTGCACAGCATTGAGGCTTACTTGCCTGA	2840
	2841	GTTGTCCGTGATCCCTGGTGTGAACGCTGCCATCTCGAG	2880
	2881	GAACTTGAGGGACGTATCTTACCGCATTCTCCTTGTACG	2920
	2921	ATGCCAGAACGTCATCAAGAACGGTACTTCAACAAATGG	2960

5	2961	CCTCAGCTGCTGGAATGTGAAAGGTATGTGGACGTGGAG	3000
10	3001	GAACAGAACAAATCAGCGTCCGTCTGGTTGCCTGAGT	3040
15	3041	GGGAAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGG	3080
20	3081	TAGAGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGA	3120
25	3121	TACGGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAA	3160
30	3161	ACACCGACGAGCTTAAGTCTCCAACGAGATCGAGAACAA	3200
35	3201	AATCTATCCAAACAACACCGTTACTTGCAACGACTACACT	3240
40	3241	GTGAATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTA	3280
45	3281	ACAGAGGTTACAACGAAGCTCCTCCGTTCCGTGACTA	3320
50	3321	TGCCTCCGTGTACGAGGGAGAAATCCTACACAGATGGCAGA	3360
55	3361	CGTGAGAACCTTGCAGTTCAACAGAGGTTACAGGGACT	3400
60	3401	ACACACCACCTCCAGTTGGCTATGTTACCAAGGAGCTTGA	3440
65	3441	GTACTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGT	3480
70	3481	GAAACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTC	3520
75	3521	TCTTGATGGAGGAA 3534 ,	

50 H. einem Struktur-Gen, das für ein insektizides Protein von *B.t.t.* codiert, mit der Sequenz:

55	1	ATGACTGCAGACAAACAACACCGAACGCCCTCGACAGTTCTA	40
	41	CCACTAAGGATGTTATCCAGAAGGGTATCTCCGTTGTGGG	80

	81	AGACCTCTGGCGTGGATTCCCTCGGTGGAGCC	120
5	121	CTCGTGAGCTCTATACAAACTTCTAACACCATTGGC	160
	161	CAAGCGAGGACCCTGGAAAGCATTCATGGAGCAAGTTGA	200
10	201	AGCTCTTATGGATCAGAAGATTGCAGATTATGCCAAGAAC	240
	241	AAGGCTTGGCAGAACTCCAGGGCCTTCAGAACATGTGG	280
15	281	AGGACTACGTGAGTGCATTGTCCAGCTGGCAGAAGAACCC	320
20	321	TGTTAGCTCCAGAAATCCTCACAGCCAAGGTAGGATCAGA	360
	361	GAGTTGTTCTCTCAAGCCGAATCCCACCTTCAGAAATTCCA	400
25	401	TGCCTAGCTTGCTATCTCCGGTTACGAGGTTCTTCCT	440
	441	CACTACCTATGCTCAAGCTGCCAACACCCACTTGTTCCTC	480
30	481	CTTAAGGACGCTCAAATCTATGGAGAAGAGTGGGGATAACG	520
	521	AGAAAGAGGACATTGCTGAGTTCTACAAGCGTCAACTAA	560
35	561	GCTCACCCAAGAGTACACTGACCATTGCGTGAATGGTAT	600
40	601	AACGTTGGTCTCGATAAGCTCAGAGGCTCTCCTACGAGT	640
	641	CTTGGGTGAACCTCAACAGATAACAGGAGAGAGATGACCTT	680
45	681	GACTGTGCTCGATCTTATCGCACTCTTCCCTGTACGAT	720
	721	GTGAGACTCTACCCAAAGGAAGTGAAAAGTCAACCAACCA	760
50	761	GAGACGTGCTCACTGACCCATTGTCGGAGTCACAAACCT	800
	801	TAGGGGTTATGGAACCTACCTTCAGCAATATCGAAAAACTAC	840

	841	ATTAGGAAACCACATCTCTCGACTATCTTCACAGAACATT	880
5	881	AATTCCACACAAGGTTCAACCAGGATACTATGGTAACGA	920
10	921	CTCCTTCAACTATTGGTCCGGTAACATATGTTCCACCAGA	960
15	961	CCAAGCATTGGATCTAATGACATCATCACATCTCCTTCT	1000
20	1001	ATGGTAACAAGTCCAGTGAACCTGTGCAGAACCTTGAGTT	1040
25	1041	CAACGGCGAGAAAGTCTATAGAGCCGTGCAAACACCAAT	1080
30	1081	CTCGCTGTGTGGCCATCCGCAGTTACTCAGGCGTCACAA	1120
35	1121	AGGTGGAGTTAGTCAGTATAACGATCAGACCGATGAGGC	1160
40	1161	CAGCACCCAGACTTACGACTCCAAACGTAACGTTGGCGCA	1200
45	1201	GTCTCTGGGATTCTATCGACCAATTGCCTCCAGAAACCA	1240
50	1241	CAGACGAACCATTGGAGAAGGGCTACAGCCACCAACTTAA	1280
55	1281	CTATGTGATGTGCTTCTTGATGCAAGGTTCCAGAGGGACC	1320
	1321	ATTCCAGTGTGACCTGGACACACACAAGTCCGTGGACTTCT	1360
	1361	TCAACATGATCGATAGCAAGAAGATCACTCAACTTCCCTT	1400
	1401	GGTGAAAGCCTACAAGCTGCAATCTGGTGCTTCCGTTGTC	1440
	1441	GCAGGTCCCAGATTCACTGGAGGTGACATCATCCAGTGCA	1480
	1481	CAGAGAACGGCAGCGCAGCTACTATCTACGTGACACCTGA	1520
	1521	TGTGTCTTACTCTCAGAAAGTACAGGGCACGTATTCAATTAC	1560
	1561	GCATCTACCAGCCAGATCACCTCACACTCAGCTTGGATG	1600

5 1601 GAGCACCCCTCAACCAGTATTACTTTGACAAGACCATCAA 1640
 . 1641 CAAAGGTGACACTCTCACATACAATAGCTCAACTGGCA 1680
 . 1681 AGTTTCAGCACACCATTGAACTCTCAGGCAACAATCTTC 1720
 . 1721 AGATCGGCGTCACCGGTCTCAGCGCCGGAGACAAAGTCTA 1760
15 1761 CATCGACAAGATTGAGTTCATCCCAGTGAAC 1791 ;

20 I. einem Struktur-Gen, das für ein insektizides Protein von *B. t. entomocidus* codiert, mit der Sequenz:

25 1 ATGGAGGAGAACAAACCAAAACCAATGCATTCCATAACACT 40
 . 41 GCTTGAGTAACCCAGAACAGAGGTATTGCTTGATGGAGAACG 80
 . 81 CATTCAACCGGTAACTCTTCATCGACATCTCCTTGTCC 120
 . 121 TTGGTCCAGTTCTGGTCAGCAACTCGTGCCAGGTGGTG 160
35 161 GGTTCTTGTCGGACTAATTGACTTCGTCTGGGTATCGT 200
 . 201 TGGTCCATCTCAATGGGATGCATTCTGGTGCAAATTGAG 240
 . 241 CAGTTGATCAACGAGAGGATCGCTGAGTTGCCAGGAACG 280
 . 281 CTGCCATCGCTAACTTGAAGGATTGGCAATAACTCAA 320
 . 321 CATCTATGTGGAGGCCTTCAAAGAGTGGGAAGAGGACCC 360
50 361 AACAAACCCAGAGACCCGCACTAGGGTATCGACAGATTCA 400
 . 401 GAATCTTGGACGGCCTCTGGAGAGAGATATCCCATCCCTT 440
 . 441 CAGAAATCTCTGGCTTCGAAGTTCCCTCTTGTCCGTGTAC 480

481	GCTCAAGCAGCTAATCTTCACCTCGCTATCCTCGAGACA	520
521	GTGTCATCTTGGGAAAGGTGGGATTGACCACTATCAA	560
561	CGTCAATGAGAATTACAACAGACTTATCAGGCACATTGAC	600
601	GAGTACGCCGACCACGTGCTAACACCTACAACCGTGGCT	640
641	TGAACAATCTCCCTAACGTCTACTTATCAAGATTGGATTAC	680
681	CTACAAACAGGTGAGGAGAGACTTGACCCCTCACAGTTTG	720
721	GACATTGCAGCTTCTTCCCAGACTATGACAACAGGAGAT	760
761	ACCCATCCAACCAGTGGGTCAACTTACCAAGAGAAGTCTA	800
801	TACTGACCCACTTATCAACTTCAACCCCTCAGTTGCAAAGT	840
841	GTGCCCAACTTCCCACATTCAACGTATGGAGTCCAGCC	880
881	GTATCAGGAACCCACACTTGTGACATCTTGAACAAACCT	920
921	TACTATCTCACCGATTGGTCAGCGTTGGCGTAACCTC	960
961	TATTGGGTGGACACAGGGTCATCTCCTCTCTTATTGGAG	1000
1001	GTGGGAACATTACCTCTCCTATCTATGGACGTGAGGAAA	1040
1041	CCAGGAGCCACCACGTAGTTCACCTCAACGGTCCAGTC	1080
1081	TTCAGAACCTTGTCTAACCTACCTTGAGATTGCTCCAGC	1120
1121	AACCTTGGCCAGCTCCACCTTCAACCTTAGAGGTGTTGA	1160
1161	GGGCGTTGAGTTCTACTCCTACCAACTCCTTCACTTAC	1200
1201	AGAGGTAGAGGAACCGTTGATTCTTGACCGAACTCCCAC	1240

5	1241	CAGAGGACAATAGCGTGCCACCCAGGGAAAGGCTACTCCCCA	1280
10	1281	CAGGGTGTGCCACGCAACCTCGTGCAGCGTCCGGAACCT	1320
15	1321	CCATTCTCACTACAGGAGTTGTGTTCTCATGGACTGATC	1360
20	1361	GTAGTGCTACTCTCACTAACCAATTGATCCCAGAGGAGAT	1400
25	1401	CAATCAAATCCCATTGGTCAAGGGTTCCGTGTGGGGAA	1440
30	1441	GGAACCTCTGTACACAGGACCAGGCTCACAGGAGGTG	1480
35	1481	ATATTCTTAGAAGAACACTTTGGCGACTTGTGAGCCT	1520
40	1521	CCAAGTTAACATCAACTCTCCAATTACTCAAAGATAATCGT	1560
45	1561	CTCAGGTTTCGTTACGCATCTTCCGTGACGCTAGAGTCA	1600
50	1601	TCGTGCTCACCGGAGCAGCTTCTACCGGTGTCGGTGGACA	1640
55	1641	AGTCTCCGTAAACATGCCACTCCAGAAGACTATGGAGATC	1680
	1681	GGCGAGAACTTGACATCCAGGACCTTCAGATAACACCGACT	1720
	1721	TCTCTAACCCTTTCAGTTCCGTGCCAACCTGACATCAT	1760
	1761	TGGCATTAGCGAACAAACCTCTTTGGAGCTGGTAGCATC	1800
	1801	TCATCTGGCGAATTGTACATTGACAAGATTGAGATCATTC	1840
	1841	TTGCCGACGCTACCTCGAGGCTGAGTCTGACCTTGAGAG	1880
	1881	AGCCCAGAAGGCTGTGAACGCCCTCTTACCTCCTCTAAT	1920
	1921	CAGATTGGCTTGAAAAGTACGTTACTGACTATCACATTG	1960
	1961	ACCAAGTGTCCAACCTGGTCGACTGCCCTAGCGATGAGTT	2000

5	2001	CTGCCTCGACCGAGAAGCGTGAACCTCTCCGAGAAAGTTAAA	2040
10	2041	CACGCCAACCGCTCTCAGCGACGAGAGGAATCTCTTGCAAG	2080
15	2081	ACCCCAACTTCAGAGGCATCAACAGGCAGCCAGACCGTG	2120
20	2121	TTGGAGAGGAAGCACCACATCACCATCCAAGGAGGCGAC	2160
25	2161	GATGTGTTCAAGGAGAACTACGTACCCCTCCCAGGAAC	2200
30	2201	TGGACGAGTGCTACCCCTACCTACTTGTACCAGAAGATCGA	2240
35	2241	TGAGTCCAAACTCAAAGCCTACACCAGGTATGAAC	2280
40	2281	GGCTACATCGAAGACAGCCAAGACCTTGAAATCTACCTCA	2320
45	2321	TCAGGTACAATGCCAAGCACGAGATCGTGAATGTCCCAGG	2360
50	2361	TACTGGTCCCTCTGGCCACTTTCTGCCAAATGCCATT	2400
55	2401	GGGAAGTGTGGAGAGCCTAACAGATGCGCTCCACACCTTG	2440
	2441	AGTGGAAATCCTGACTTGGACTGCTCTGCAGGGATGGCGA	2480
	2481	GAAGTGTGCCACCATTCTCATCACCTCACCTGGACATC	2520
	2521	GATGTGGGATGTACTGACCTGAATGAGGACCTCGGAGTCT	2560
	2561	GGGTCATCTCAAGATCAAGACCCAAAGACGGACACGCAAG	2600
	2601	ACTTGGCAACCTTGAGTTCTCGAAGAGAAACCATTGCTC	2640
	2641	GGTGAAGCTCTCGCTCGTGTGAAGAGAGCAGAGAAGAAGT	2680
	2681	GGAGGGACAAACGTGAGAAACTCCAACTCGAGACTAACAT	2720
	2721	CGTTTACAAGGAGGCCAAAGAGTCCGTGGATGCTTGTTC	2760

5	2761	GTGAACTCCCAATATGATAGGTTGCAAGTGGACACCAACA	2800
10	2801	TCGCCATGATCCACGCTGCAGACAAACGTGTGCACAGGAT	2840
15	2841	TCGTGAGGCTTACTTGCTGAGTTGTCCGTGATCCCTGGT	2880
20	2881	GTGAACGCTGCCATCTCGAGGAACTTGAGGGACGTATCT	2920
25	2921	TTACCGCATACTCCTTGTACGATGCCAGAAACGTCAAA	2960
30	2961	GAACGGTGACTTCAACAATGGCCTCTTGTGCTGGAATGTG	3000
35	3001	AAAGGTCATGTGGACGTGGAGGAACAGAACAAATCACCGTT	3040
40	3041	CCGTCCCTGGTTATCCCTGAGTGGGAAGCTGAAGTGTCCCCA	3080
45	3081	AGAGGTTAGAGTCTGTCCAGGTAGAGGCTACATTCTCCGT	3120
50	3121	GTGACCGCTTACAAGGAGGGATACGGTGAGGGTTGCGTGA	3160
55	3161	CCATCCACGAGATCGAGGACAACACCGACGAGCTTAAGTT	3200
60	3201	CTCCAAC TGCGTCGAGGAAGAAGTCTATCCCAACAAACACC	3240
65	3241	GTTACTTGCAACAACTACACTGGGACCCAGGAAGAGTACG	3280
70	3281	AAGGTACCTACACTAGCCGTAACCAAGGTTACGACGAAGC	3320
75	3321	TTACGGAAACAATCCTTCCGTTCTGCTGACTATGCCTCC	3360
80	3361	GTGTACGAGGAGAAATCCTACACAGATGGCAGACGTGAGA	3400
85	3401	ACCCCTTGCAGTCCAACAGAGGTTACGGTGAACACACCC	3440
90	3441	ACTTCCAGCAGGCTATGTTACCAAGGACCTTGAGTACTTT	3480
95	3481	CCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAAACCG	3520

5 3521 AGGGAACCTTCATCGTGGACAGCGTGGAGCTCTTGAT 3560

10 3561 GGAGGAA 3567.

15 J. einem Struktur-Gen, das für ein insektizides P2-Protein codiert, mit der Sequenz:

15	1 ATGGACAACAACGTCTGAACCTCTGGTAGAACAAACCACATCT	40
	41 GCGACGCATAACAACGTCTGGCTACGATCCATTCAAGCTT	80
20	81 CGAACACAAGAGCCTCGACACTATTCAAGAAGGAGTGGATG	120
	121 GAATGGAAACGTACTGACCACTCTCTACGTGCGCACCTG	160
25	161 TGGTTGGAACAGTGTCCAGCTTCCTCTCAAGAAGGTCGG	200
30	201 CTCTCTCATCGAAAACGTATCTTGTCCGAACCTGGGGT	240
	241 ATCATCTTCCATCTGGGTCCACTAATCTCATGCAAGACA	280
35	281 TCTTGAGGGAGACCGAACAGTTCTCAACCAGCGTCTCAA	320
40	321 CACTGATAACCTTGGCTAGAGTCAACGCTGAGTTGATCGGT	360
	361 CTCCAAGCAAACATTGTGAGTTCAACCAGCAAGTGGACA	400
45	401 ACTTCTTGAATCCAACCTCAGAACATCTGTGCCCTTTCCAT	440
	441 CACTTCTTCCGTGAACACTATGCAGCAACTCTTCCTCAAAC	480
50	481 AGATTGCCTCAGTTCAAGGCTACCAGTTGCTCC	520
	521 TTCTTCCACTCTTGTCAAGGCTGCCAACATGCACTTGTC	560
55	561 CTTCATACGTGACGTGATCCTCAACGCTGACGAATGGGGA	600

5	601	ATCTCTGCAGCCACTCTTAGGACATACTAGAGACTACTTGA	640
10	641	GGAACACTACACTCGTGATTACTCCAACATTGCATCAACAC	680
15	681	TTATCAGACTGCCCTTCGTGGACTCAATACTAGGCTTCAC	720
20	721	GACATGCTTGAGTTCAAGGACCTACATGTTCTAACGTGT	760
25	761	TTGAGTACGTCAGCATTTGGAGTCTCTCAAGTACCGAGAG	800
30	801	CTTGATGGTGTCTCTGGAGCCAATCTCTACGCCCTCTGGC	840
35	841	AGTGGACCACAGCAAACCTCAGAGCTTCACAGCTCAGAACT	880
40	881	GGCCATTCTTGTATAGCTTGTCCAAGTCAACTCCAACTA	920
45	921	CATTCTCAGTGGTATCTCTGGGACCAGACTCTCCATAACC	960
50	961	TTTCCCAACATTGGTGGACTTCCAGGCTCCACTACAACCC	1000
55	1001	ATAGCCTTAACCTCTGCCAGAGTGAACACTACAGTGGAGGTGT	1040
60	1041	CAGCTCTGGATTGATTGGTGCAACTAACTTGAACCACAAC	1080
65	1081	TTCAATTGCTCCACCGTCTGCCACCTCTGAGCACACCGT	1120
70	1121	TTGTGAGGTCTGGCTTGACAGCGGTACTGATCGCGAAGG	1160
75	1161	AGTTGCTACCTCTACAAACTGGCAAACCGAGTCCTTCAA	1200
80	1201	ACCACTCTTAGCCTCGGTGTGGAGCTTCTCTGCACGTG	1240
85	1241	GGAATTCAAACACTTTCCAGACTACTTCATTAGGAACAT	1280
90	1281	CTCTGGTGTCTCGTCATCAGGAATGAAGACCTCACC	1320
95	1321	CGTCCACTTCATTACAACCAGATTAGGAACATCGAGTCTC	1360

5 1361 CATCCGGTACTCCAGGAGGTGCAAGAGCTTACCTCGTGTC 1400

10 1401 TGTCCATAACAGGAAGAACAAACATCTACGCTGCCAACGAG 1440

15 1441 AATGGCACCATGATTCACCTTGCACCAGAAGATTACACTG 1480

20 1481 GATTCAACCCTCTCCAATCCATGCTACCCAAGTGAACAA 1520

25 1521 TCAGACACGCACCTTCATCTCCGAAAAGTTCGGAAATCAA 1560

30 1561 GGTGACTCCTTGAGGTTCGAGCAATCCAACACTACCGCTA 1600

35 1601 GGTACACTTGAGAGGCAATGGAAACAGCTACAACCTTTA 1640

40 1641 CTTGAGAGTTAGCTCCATTGGTAACCTCACCACCGTGTGTT 1680

45 1681 ACCATCAACGGACGTGTTACACAGTCTCTAATGTGAACA 1720

50 1721 CTACAACGAACAATGATGGCGTTAACGACAACGGAGCCAG 1760

55 1761 ATTCAAGCGACATCAACATTGGCAACATCGTGGCCTCTGAC 1800

60 1801 AACACTAACGTTACTTGACATCAATGTGACCCCTCAATT 1840

65 1841 CTGGAACTCCATTGATCTCATGAACATCATGTTGTGCC 1880

70 1881 AACTAACCTCCCTCCATTGTAC 1902

oder

50 K. einer Struktur-Gen-Sequenz, die für ein Fusionsprotein codiert, das die N-terminalen 610 Aminosäuren von B.t.k. HD-1 und die C-terminalen 567 Aminosäuren von B.t.k. HD-73 aufweist, welches Gen die Sequenz hat:

55 1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40

41	ACTGCTTGAGTAACCCAGAAGTGAAGTACTTGGTGGAGA	80
5		
81	ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG	120
10		
121	TCCTTGACACAGTTCTGCTCAGCGAGTTCGTGCAGGTG	160
15		
161	CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGTAT	200
20		
201	CTTTGGTCCATCTCAATGGGATGCATTCTGGTGCAAATT	240
25		
241	GAGCAGTTGATCAACCAGAGGATCGAAGAGTTGCCAGGA	280
30		
281	ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA	320
35		
321	CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT	360
40		
361	CCTACTAACCCAGCTCTCCGGAGGAAATGCGTATTCAAT	400
45		
401	TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT	440
50		
441	GTTCGCAGTCCAGAACTACCAAGTTCCCTCTTGTCCGTG	480
55		
481	TACGTTCAAGCAGCTAACCTCACCTCAGCGTGCCTCGAG	520
60		
521	ACGTTAGCGTGTGGGCAAAGGTGGGATTGATGCTGC	560
65		
561	AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT	600
70		
601	GGAAACTACACCGACCACGCTGTTGTTGGTACAACACTG	640
75		
641	GCTTGGAGCGTGTCTGGGTCTGATTCTAGAGATTGGAT	680
80		
681	TAGATAACAACCAGTTCAGGAGAGAATTGACCCCTCACAGTT	720
85		
721	TTGGACATTGTGTCTCTTCCGAACTATGACTCCAGAA	760
90		
761	CCTACCCCTATCCGTACAGTGTCCCAACTTACCAAGAGAAAT	800

5	801	CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC	840
	.	.	.
10	841	CGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
	.	.	.
15	881	CACACTTGATGGACATCTGAACAGCATAACTATCTACAC	920
	.	.	.
20	921	CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG	960
	.	.	.
25	961	ATCATGGCCTCTCCAGTTGGATTCAAGCGGGCCCGAGTTA	1000
	.	.	.
30	1001	CCTTCCTCTATGAACTATGGAAACGCCGCTCCACA	1040
	.	.	.
35	1041	ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA	1080
	.	.	.
40	1081	ACCTTGTCTTCCACCTTGTACAGAAGACCCCTCAATATCG	1120
	.	.	.
45	1121	GTATCAACAACCAGCAACTTCCGTTCTGACGGAACAGA	1160
	.	.	.
50	1161	GTTCGCCTATGAAACCTCTTCTAACTTGCCATCCGCTGTT	1200
	.	.	.
55	1201	TACAGAAAGAGCGGAACCGTTGATTCCCTGGACGAAATCC	1240
	.	.	.
60	1241	CACCACAGAACACAATGTGCCACCCAGGCAAGGATTCTC	1280
	.	.	.
65	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC	1320
	.	.	.
70	1321	AGCAACAGTTCCGTGAGGCATCATCAGAGCTCCTATGTTCT	1360
	.	.	.
75	1361	CATGGATTCATCGTAGTGCTGAGTTCAACAATATCATTCC	1400
	.	.	.
80	1401	TTCCTCTCAAATCACCCAAATCCCATTGACCAAGTCTACT	1440
	.	.	.
85	1441	AACCTTGGATCTGGAACCTTCTGTGAGGCTAAAGGACCAGGCT	1480
	.	.	.
90	1481	TCACAGGAGGTGATATTCTTAGAAGAACTTCTCCTGGCCA	1520

	1521	GATTAGCACCCCTCAGAGTTAACATCACTGCACCACCTTCT	1560
5		.	.
	1561	CAAAGATATCGTGTCAAGGATTCGTACGCATCTACCACTA	1600
10		.	.
	1601	ACTTGCAATTCCACACCTCCATCGACGGAAGGCCTATCAA	1640
15		.	.
	1641	TCAGGGTAACCTCTCCGCAACCATGTCAAGCGGCAGCAAC	1680
20		.	.
	1681	TTGCAATCCGGCAGCTTCAGAACCGTCGGTTCACTACTC	1720
25		.	.
	1721	CTTCAACTTCTCTAACGGATCAAGCGTTTCACCCTTAG	1760
30		.	.
	1761	CGCTCATGTGTTCAATTCTGGCAATGAAGTGTACATTGAC	1800
35		.	.
	1801	CGTATTGAGTTGTGCCTGCCGAAGTTACCCTCGAGGCTG	1840
40		.	.
	1841	AGTACAACCTTGAGAGAGGCCAGAAGGCTGTGAACGCCCT	1880
45		.	.
	1881	CTTACCTCCACCAATCAGCTGGCTTGAAACTAACGTT	1920
50		.	.
	1921	ACTGACTATCACATTGACCAAGTGTCCAACCTGGTCACCT	1960
55		.	.
	1961	ACCTTAGCGATGAGTTCTGCCTCGACGAGAAGCGTGAAC	2000
		.	.
	2001	CTCCGAGAAAGTTAACACGCCAAGCGTCTCAGCGACGAG	2040
60		.	.
	2041	AGGAATCTTGTCAAGACTCCAACCTCAAAGACATCAACA	2080
65		.	.
	2081	GGCAGCCAGAACGTGGTGGGATCAC	2120
70		.	.
	2121	CATCCAAGGAGGCAGCATGTGTTCAAGGAGAACTACGTC	2160
75		.	.
	2161	ACCCTCTCCGGAACCTTCGACGAGTGCTACCTACCTACT	2200
80		.	.
	2201	TGTACCAAGAGATCGATGAGTCCAACCTCAAAGCCTTCAC	2240
85		.	.
	2241	CAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAAGAC	2280

5	2281	CTTGAATCTACTCGATCAGGTACAATGCCAAGCACGAGA	2320
10	2321	CCGTGAATGTCCCAGGTACTGGTCCCTCTGGCCACTTTC	2360
15	2361	TGCCCAATCTCCCATTGGGAAGTGTGGAGAGCCTAACAGA	2400
20	2401	TGCGCTCCACACCTTGAGTGGAAATCCTGACTGGACTGCT	2440
25	2441	CCTGCAGGGATGGCGAGAAGTGTGCCACCATTCTCATCA	2480
30	2481	CTTCTCCTTGGACATCGATGTGGATGTACTGACCTGAAT	2520
35	2521	GAGGACCTCGGAGTCTGGTCATCTCAAGATCAAGACCC	2560
40	2561	AAGACGGACACGCAAGACTTGGCAACCTTGAGTTCTCGA	2600
45	2601	AGAGAAACCATTGGTCGGTGAAGCTCTCGCTCGTGTGAAG	2640
50	2641	AGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAAACTCG	2680
55	2681	AATGGGAAACTAACATCGTTACAAGGAGGCCAAAGAGTC	2720
	2721	CGTGGATGCTTGTTCGTGAACCTCCAATATGATCAGTTG	2760
	2761	CAAGCCGACACCAACATGCCATGATCCACGCCGCAGACA	2800
	2801	AACGTGTGCACAGCATTCTGAGGGTTACTTGCCTGAGTT	2840
	2841	GTCCGTGATCCCTGGTGTGAACGCTGCCATCTCGAGGAA	2880
	2881	CTTGAGGGACGTATCTTACCGCATTCTCCTGTACGATG	2920
	2921	CCAGAAACGTCAAGAACGGTGACTTCAACAATGGCCT	2960
	2961	CAGCTGCTGGAATGTGAAAGGTCAATGTGGACGTGGAGGAA	3000
	3001	CAGAACAAATCAGCGTCCGTCTGGTTGTGCCTGAGTGGG	3040

5	3041	AAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGGTAG	3080
10	3081	AGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGATAC	3120
15	3121	GGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAAACA	3160
20	3161	CCGACGAGCTTAAGTTCTCCAAC TGCGTCGAGGAAGAAAT	3200
25	3201	CTATCCAAACAACACCGTTACTTGCAACGACTACACTGTG	3240
30	3241	AATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTAACA	3280
35	3281	GAGGTTACAACGAAGCTCCTTCCGTTCTGCTGACTATGC	3320
40	3321	CTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGACGT	3360
45	3361	GAGAACCC TTGCGAGTTAACAGAGGTTACAGGGACTACA	3400
50	3401	CACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGAGTA	3440
55	3441	CTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAA	3480
60	3481	ACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTCTCT	3520
65	3521	TGATGGAGGAA 3531.	

Revendications

1. Procédé de modification d'une séquence de gène de structure du type sauvage qui code une protéine insecticide de *Bacillus thuringiensis* afin d'activer l'expression de ladite protéine chez des plantes qui comprend :
 - a) l'identification de régions à l'intérieur de ladite séquence comprenant plus de quatre nucléotides consécutifs d'adénine ou de thymine,
 - b) la modification des régions de l'étape a) qui comportent deux ou plusieurs signaux de polyadénylation à l'intérieur d'une séquence de dix bases afin d'éliminer lesdits signaux tout en conservant une séquence de gène qui code ladite protéine, et
 - c) la modification des régions de 15 à 30 bases entourant les régions de l'étape a) afin d'éliminer les signaux majeurs de polyadénylation de plantes, les séquences consécutives contenant plus d'un signal mineur de polyadénylation et les séquences consécutives contenant plus d'une séquence ATTAA tout en conservant une séquence de gène qui code ladite protéine.

2. Procédé de modification d'une séquence de gène de structure du type sauvage qui code une protéine insecticide de *Bacillus thuringiensis* afin d'activer l'expression de ladite protéine chez des plantes qui comprend :
- a) l'élimination des signaux de polyadénylation contenus dans ledit gène de type sauvage tout en conservant une séquence qui code ladite protéine, et
 b) l'élimination des séquences ATTTA contenues dans ledit gène de type sauvage tout en conservant une séquence qui code ladite protéine.
3. Procédé selon la revendication 2, comprenant en outre l'élimination des séquences autocomplémentaires et le remplacement de telles séquences par de l'ADN non autocomplémentaire comprenant des codons préférés des plantes tout en conservant une séquence de gène de structure codant ladite protéine.
4. Procédé selon les revendications 1 à 3, comprenant en outre l'utilisation des séquences préférées des plantes au cours de l'élimination des signaux de polyadénylation et des séquences ATTTA.
5. Procédé selon les revendications 1 à 3, dans lequel les signaux de polyadénylation des plantes sont choisis parmi le groupe constitué de AATAAA, AATAAT, AACCAA, ATATAA, AATCAA, ATACTA, ATAAAA, ATGAAA, AAGCAT, ATTAAT, ATACAT, AAAATA, ATTAAA, AATTAA, AATACA et CATAAA.
6. Procédé destiné à améliorer l'expression d'un gène hétérologue chez des plantes dans lequel ledit gène comprend un gène chimère modifié comprenant un promoteur qui agit dans les cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont une partie au moins est dérivée d'une protéine de *Bacillus thuringiensis*, dans lequel ledit procédé comprend la modification de ladite séquence de structure codante de sorte que ladite séquence comporte une séquence d'ADN qui diffère de la séquence d'ADN apparaissant dans la nature codant ladite protéine de *Bacillus thuringiensis* et ladite séquence de structure codante ne contient pas plus de 5 nucléotides consécutifs constitués de restes soit adénine, soit thymine.
7. Procédé d'amélioration de l'expression d'un gène hétérologue chez des plantes dans lequel ledit gène comprend un gène chimère modifié comprenant un promoteur qui agit dans des cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée d'une protéine de *Bacillus thuringiensis*, dans lequel ledit procédé comprend la modification de ladite séquence de structure codante de sorte que ladite séquence comporte une séquence d'ADN qui diffère de la séquence d'ADN qui apparaît dans la nature codant ladite protéine de *Bacillus thuringiensis* et présente les caractéristiques suivantes :
- 40 ladite séquence de structure codante comporte une région qui est complémentaire de la séquence suivante :

GGCTTGATTCTCTAGCGAACTCTCGATTCTCTGGTTGATGAGCTGTTC

1 5 10 15 20 25 30 35 40 45

- 45 ladite région dans ladite séquence codante ayant éliminé 2 séquences AACCAA et 1 séquence AATTAA.
- 50 8. Procédé selon la revendication 7, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée de *Bacillus thuringiensis kurstakis HD-1*.
- 55 9. Procédé selon la revendication 7 ou 8, dans lequel la plante est un plan de tabac.
10. Gène chimère modifié contenant un promoteur qui agit dans des cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée d'une

protéine de *Bacillus thuringiensis*, dans lequel ladite séquence de structure codante comporte une séquence d'ADN qui diffère de la séquence d'ADN apparaissant dans la nature codant ladite protéine de *Bacillus thuringiensis* et est choisie à partir de :

- 5 A. Un gène de structure qui code une protéine insecticide de *B.t.k.* HD-1 comportant la séquence :

10

15

20

25

30

35

40

45

50

55

5	1	ATGGCTATAGAAA 5	40
10	41	TGTCGCTAACGCAATT 10	80
15	81	TGCTGGATTGTGTTAGGACTAG 15	120
20	121	ATTTTGGTCCCTCTCAATGGGACGC 20	160
25	161	TTGAACAGCTCATCAACCAGAGA 25	200
30	201	GAATCAAGCCATTCTAGATTAGA 30	240
35	241	TATCAAATTACGCAGAA 35	280
40	281	ATCCTACTAATCCAGCATTAAAGAGA 40	320
45	321	ATTCAATGACATGAACAGTGCCCT 45	360
50	361	CTTTTGCAGTCAAAATTATCAAG 50	400
55	401	TGTACGTTCAAGCTGCCAAC 55	440
60	441	AGATGTTTCAGTGGACAAAGGT 60	480
65	481	GCGACTATCAATAGTCGTTATA 65	520
70	521	TTGGCAACTATACAGATCAT 70	560
75	561	GGGATTAGAGCGTGTAT 75	600
80	601	ATCAGGTACAACCAGTT 80	640
85	641	TATTAGATATCGTTCTCT 85	680
90	681	AACGTATCCAATT 90	720

5	721	ATTTATAACAAACCCAGTATTAGAAAATTTGATGGTAGTT	760
	.	.	.
10	761	TTCGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAG	800
	.	.	.
15	801	TCCACATTTGATGGATATACTTAATAGTATAACCATCTAT	840
	.	.	.
20	841	ACGGATGCTCATAGAGGAGAATACTACTGGTCCGGTCACC	880
	.	.	.
25	881	AGATCATGGCTTCTCCTGTAGGGTTTCGGGGCCAGAATT	920
	.	.	.
30	921	CACTTTCCGCTATATGGAACATATGGAAATGCAGCTCCA	960
	.	.	.
35	961	CAACAAACGTATTGTTGCTCAACTAGGTCAAGGGCGTGTATA	1000
	.	.	.
40	1001	GAACATTATCGTCCACCTTATATAGAACGACCTTTAACAT	1040
	.	.	.
45	1041	CGGGATCAACAACCAACAACTATCTGTTCTTGACGGGACA	1080
	.	.	.
50	1081	GAATTGCTTATGGAACCTCCTCAAATTGCCATCCGCTG	1120
	.	.	.
55			

5	1121	TATACAGRAAAAGCGGAACGGTAGATTGCTGGATGAAAT	1160
10	1161	ACCGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTT	1200
15	1201	AGTCATCGATTAAGCCATGTTCAATGTTCGTCAGGCT	1240
20	1241	TTAGTAATAGTAGTGTAAAGTATAATAAGAGCTCCTATGTT	1280
25	1281	CTCTTGGATACATCGTAGTGCTGAGTTCAACAAACATCATC	1320
30	1321	CCTTCATCACAAATCACCCAAATCCCACTCACCAAGTCTA	1360
35	1361	CTAATCTTGGCTCTGGAACCTCTGTCGTTAAAGGACCAGG	1400
40	1401	ATTTACAGGAGGAGATATTCTCGAAGAACTCACCTGGC	1440
45	1441	CAGATTCAACCTTAAGAGTAAATATTACTGCACCATTAT	1480
50	1481	CACAAAGATATCGGGTAAGAATTGCTACGCTTCTACCAC	1520
55	1521	AAACCTTCAGTTCCACACATCAATTGACGGAAGACCTATT	1560
60	1561	AATCAGGGGAATTTTCAGCAACTATGAGTAGTAGGGAGTA	1600
65	1601	ATTTACAGTCCGGAAGCTTCTGGACTGTAGGTTTACTAC	1640
70	1641	TCCGTTAACCTTCAAATGGATCAAGTGTATTTACGTTA	1680
75	1681	AGTGCTCATGTCTCAATTCAAGGCAATGAAGTTATATAG	1720
80	1721	ATCGAATTGAATTGTTCCGGCA 1743.	

B. Un gène de structure qui code une protéine insecticide de *B.t.k.* HD-73 comportant la séquence :

5	1	ATGGCCATTGAAACCGGTTACACTCCCATCGACATCTCCT	40
10	41	TGTCCCTTGACACAGTTCTGCTCAGCGAGTTCTGCCAGG	80
15	81	TGCTGGGTTCTCGGACTAGTTGACATCATCTGGGGT	120
20	121	ATCTTGGTCCATCTCAATGGGATGCATTCTGGTGCAA	160
25	161	TTGAGCAGTTGATCAACCAGAGGATCGAAGAGTTGCCAG	200
30	201	GAACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTC	240
35	241	TACCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCG	280
40	281	ATCCTACTAACCCAGCTCTCGCGAGGAAATGCGTATTCA	320
45	321	ATTCAACGACATGAACAGCGCCTTGACCACAGCTATCCCA	360
50	361	TTGTTCGCAGTCCAGAACTACCAAGTTCTCTCTTGTCAG	400
55	401	TGTACGTTCAAGCAGCTAACTTACCTCACCTCAGCGTGCTTCG	440

5	441	AGACGTTAGCGTGTGGCAAAGGTGGGATTGATGCT	480
	481	GCAACCATAATGCCGTTACAACGACCTTAAGGCTGA	520
10	521	TTGGAAACTACACCGACCACGCTGTTGGTACAACAC	560
	561	TGGCTTGGAGCGTGTCTGGGTCTGATTCTAGAGATTGG	600
15	601	ATTAGATAACAACCAGTTCAAGGAGAGAATTGACCTCACAG	640
20	641	TTTTGGACATTGTGTCTCTCTCCGAACATGACTCCAG	680
	681	AACCTACCCATCCGTACAGTGTCCCAACTTACCAAGAGAA	720
25	721	ATCTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCT	760
	761	TCCGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAG	800
30	801	CCCACACTTGTGGACATCTTGAACAGCATAACTATCTAC	840
	841	ACCGATGCTCACAGAGGAGTATTACTGGCTGGACACC	880
35	881	AGATCATGGCCTCTCCAGTTGGATTCAAGCGGGCCCGAGTT	920
40	921	TACCTTTCTCTATGAACTATGGAAACCGCCGCTCCA	960
	961	CAACAAACGTATCGTTGCTCAACTAGGTCAAGGGTGTCAACA	1000
45	1001	GAACCTTGTCTCCACCTTGTACAGAAGACCCCTCAATAT	1040

50

55

5	1041	CGGTATCAACAACCAGCAACTTCCGTTCTTGACGGAACA	1080
10	1081	GAGTCGCCTATGGAACCTCTCTAACCTGCCATCCGCTG	1120
15	1121	TTTACAGAAAGAGCGGAACCGTTGATTCCCTGGACGAAAT	1160
20	1161	CCCACACAGAACAAACAAATGTGCCACCCAGGCAAGGATT	1200
25	1201	TCCCCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGAT	1240
30	1241	TCAGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTT	1280
35	1281	CTCTGGATAACACCGTAGTGCTGAGTTCAACAAACATCATC	1320
40	1321	GCATCCGATAAGTATTACTCAAATCCCTGCAGTGAAGGGAA	1360
45	1361	ACTTTCTCTTCAACGGTTCTGTCAATTTCAGGACCAGGATT	1400
50	1401	CACTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAAT	1440
55	1441	AACATTCAAGAATAGAGGGTATATTGAAGTTCCAATTCACT	1480
	1481	TCCCCATCCACATCTACCAGATATAGAGTTCTGTGAGGTA	1520
	1521	TGCTTCTGTGACCCCTATTCACCTAACGTTAATTGGGGT	1560
	1561	AATTCACTCCATCTCTCCAAATACAGTTCCAGCTACAGCTA	1600
	1601	CCTCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTT	1640

5 1641 TGAAAGTGCCAATGCTTTACATCTTCACTCGGTAACATC 1680
 .
 .
 .
 10 1681 GTGGGTGTTAGAAACTTACTGTCAGGAGTGATTA 1720
 .
 .
 1721 TCGACAGATTGAGTTCATCCAGTTACTGCAACACTCGA 1760
 .
 .
 15 1761 GGCTGAG 1767.

C. Un gène de structure codant une protéine insecticide de *B.t.k.* HD-1 comportant la séquence :

20
 .
 .
 .
 1 ATGGACAACAACCAAACATCAACGAATGCATTCCATACA 40
 .
 .
 25 41 ACTGCCTGAGTAACCCAGAAGTGAAGTACTTGGTGGAGA 80
 .
 .
 81 ACGCATTGAAACCAGGTTACACTCCCATCGACATCTCCTTG 120
 .
 .
 30 121 TCCTTGACACAGTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
 .
 .
 35 161 CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGTAT 200
 .
 .
 201 CTTTGGTCCATCTCAATGGGATGCATTCTGGTGCAAATT 240
 .
 .
 40 241 GAGCAGTTGATCAACCAGAGGGATCGAAGAGTTGCCAGGA 280
 .
 .
 281 ACCAGGCCATCTCTAGGTTGGAAGGGATTGAGCAATCTCTA 320
 .
 .
 45 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGAAAAGCCGAT 360

50

55

5 361 CCTACTAACCCAGCTCTCCGGAGGAAATGGTATTCAAT 400
 10 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT 440
 15 441 GTTCGCAGTCCAGAACTACCAAGTTCTCTTGTCCGTG 480
 20 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGTTCGAG 520
 25 521 ACGTTAGCGTGTGGGCAAAGGTGGGGATTGATGCTGC 560
 30 561 AACCATCAATAAGCCGTTACAACGACCTTACTAGGCTGATT 600
 35 601 GGAAACTACACCGACCACGCTGTCGTTGGTACAACACTG 640
 40 641 GCTTGGAGCGTGTCTGGGGTCTGATTCTAGAGATTGGAT 680
 45 681 TAGATACAACCAGTTCAGGAGAGAATTGACCCTCACAGTT 720
 50 721 TTGGACATTGTGTCTCTTCCCCAACTATGACTCCAGAA 760
 55 761 CCTACCCCTATCCGTACAGTGTCCCAACTTACCAAGAGAAAT 800
 60 801 CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC 840
 65 841 CGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAGCC 880
 70 881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920
 75 921 CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCCAG 960

5	961	ATCATGGCCTCTCCAGTTGGATTCAAGCGGGCCCCGAGTTTA	1000
10	1001	CCTTTCCCTCTCTATGGAACATATGGAAACGCCGCTCCACA	1040
15	1041	ACAAACGTATCGTTGCTCAACTAGGTCAAGGGTGTCTACAGA	1080
20	1081	ACCTTGTCTTCCACCTTGTACAGAACGACCCCTCAATATCG	1120
25	1121	GTATCAACAAACCAGCAACTTTCCGTTCTTGACGGAACAGA	1160
30	1161	GTTGCCTATGGAACCTCTTCTAACCTGCCATCCGCTGTT	1200
35	1201	TACAGAAAGAGCGGAACCGTTGATTCCGGACGAAATCC	1240
40	1241	CACCAACAGAACAAACAAATGTGCCACCCAGGCAAGGATTCTC	1280
45	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTC	1320
50	1321	AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT	1360
55	1361	CATGGATTCACTCGTAGTGCTGAGTTCAACAAATATCATTCC	1400
	1401	TTCCCTCTCAAATCACCCAAATCCCATTGACCAAGTCTACT	1440
	1441	AACCTTGGATCTGGAACCTCTGTGCGTGAAGGACCAGGCT	1480
	1481	TCACAGGGAGGTGATATTCTTAGAAGAACCTTCTCCTGGCCA	1520
	1521	GATTAGCACCCCTCAGAGTTAACATCACTGCACCACTTTCT	1560

5 1561 CAAAGATACTGTCAGGATTCTTACGCATCTACCACTA 1600
 .
 .
 .
 10 1601 ACTTGCAATTCCACACCTCCATCGACGGAAGGCCTATCAA 1640
 .
 .
 .
 1641 TCAGGGTAACCTCTCCGAACCATGTCAAGCGGCAGCAAC 1680
 .
 .
 .
 1681 TTGCAATCCGGCAGCTTCAGAACCGTCGGTTCACTACTC 1720
 .
 .
 .
 15 1721 CTTTCAACTCTCTAACGGATCAAGCGTTTACCCCTTAG 1760
 .
 .
 .
 20 1761 CGCTCATGTGTTCAATTCTGGCAATGAAGTGTACATTGAC 1800
 .
 .
 .
 1801 CGTATTGAGTTGTGCCCTGCCGAAGTTACCTTCGAGGCTG 1840
 .
 .
 .
 25 1841 AGTAC 1845.

D. Un gène de structure codant une protéine insecticide dérivée de *B.t.k.* HD-73 comportant la séquence :

30 1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40
 .
 .
 .
 35 41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA 80
 .
 .
 .
 40 81 ACGCATTTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120
 .
 .
 .
 45 121 TCCTTGACACAGTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
 .
 .
 .
 161 CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGTAT 200

5	201	CTTTGGTCCATCTCAATGGGATGCATTCTGGTCAAATT	240
10	241	GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA	280
15	281	ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA	320
20	321	CCAAATCTATGCAGAGAGCTTCAGAGAGTGAAAAGCCGAT	360
25	361	CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT	400
30	401	TCAACGGACATGAACAGCGCCTTGACCACAGCTATCCCATT	440
35	441	GTTCGCAGTCCAGAACTACCAAGTTCTCTCTGTCCGTG	480
40	481	TACGTTCAAGCAGCTAATCTCACCTCAGCGTGCTTCGAG	520
45	521	ACGTTAGCGTGTGGGAAAGGTGGGATTGATGCTGC	560
50	561	AACCATAATAGCCGTTACAACGACCTTACTAGGCTGATT	600
55	601	GGAAACTACACCGACCACCGCTGTTGGTACAACACTG	640
60	641	GCTTGGAGCGTGTCTGGGGCTGATTCTAGAGATTGGAT	680
65	681	TAGATAACAACCAGTTCAAGGAGAGAATTGACCCCTCACAGTT	720
70	721	TTGGACATTGTGTCTCTTCCCGAACTATGACTCCAGAA	760
75	761	CCTACCCTATCCGTACAGTGTCCCAACTTACCAAGAGAAAT	800

5	801	CTATACTAACCCAGTTCTTGAGAAGCTTCGACGGTAGCTTC	840
		.	.
	841	CGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
10		.	.
	881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
		.	.
15	921	CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG	960
		.	.
20	961	ATCATGGCCTCTCCAGTTGGATTCAAGCGGGCCCGAGTTTA	1000
		.	.
	1001	CCTTCCTCTATGGAACATATGGAAACGCCGCTCCACA	1040
25		.	.
	1041	ACAAACGTATCGTTGCTCAACTAGGTCAAGGGTGTCTACAGA	1080
		.	.
	1081	ACCTTGTCTTCCACCTTGTACAGAAGACCCTCAATATCG	1120
30		.	.
	1121	GTATCAACAAACCAGCAACTTCCGTTCTTGACGGAACAGA	1160
		.	.
35	1161	GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT	1200
		.	.
	1201	TACAGAAAGAGCGGAACCGTTGATTCCCTGGACGAAATCC	1240
40		.	.
	1241	CACCAACAGAACAAACATGTGCCACCCAGGCAAGGATTCTC	1280
		.	.
	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTG	1320
45		.	.
	1321	AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT	1360
		.	.
50	1361	CTTGGATAACACCGTAGTGCTGAGTTCAACAAACATCATCGC	1400

5	1401	ATCCGATACTTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
10	1441	TTTCTCTCAACGGTTCTGTCAATTTCAGGACCAGGATTCA	1480
15	1481	CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA	1520
20	1521	CATTAGAATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
25	1561	CCATCCACATCTACCAGATATAGAGTTCTGTGTGAGGTATG	1600
30	1601	CTTCTGTGACCCCTATTCACCTAACGTTAATTGGGGTAA	1640
35	1641	TTCATCCATCTTCTCCAATACAGTTCCAGCTACAGCTACC	1680
40	1681	TCCCTGGATAATCTCCAATCCAGCGATTCGGTTACTTTG	1720
45	1721	AAAGTGCCAATGCTTTACATCTTCACTCGGTAACATCGT	1760
50	1761	GGGTGTTAGAAACTTTAGTGGACTGCAGGAGTGATTATC	1800
55	1801	GACAGATTGAGTTCATCCAGTTACTGCAACACTCGAGG	1840
60	1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCCGTAATGCG	1880
65	1881	CTGTTACGTCTACAAACCAGCTTGGACTCAAGACAAATG	1920
70	1921	G 1921;	

E. Un gène de structure codant la protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

1 ATGGACAACAAACCCAAACATCAACGAATGCATTCCATACA 40
 5

41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA 80
 10

81 ACGCATTGAAACCGGTACACTCCCATCGACATCTCCTTIG 120
 15

121 TCCTTGACACAGTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
 20

161 CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGTAT 200
 25

201 CTTTGGTCCATCTCAATGGGATGCATTCTGGTGCAAATT 240
 30

241 GAGCAGTTGATCAACCAGAGGATCGAAGAGAGTCGCCAGGA 280
 35

281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320
 40

321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360
 45

361 CCTACTAACCCAGCTCTCCGCGAGGAAATGGTATTCAAT 400
 50

401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT 440
 55

441 GTTCGCAGTCCAGAACTACCAAGTTCTCTCTGTCCGTG 480
 60

481 TACGTTCAAGCAGCTAATCTCACCTCAGCGTGCTTCGAG 520
 65

521 ACGTTAGCGTGTGGCAAAGGTGGGGATTGATGCTGC 560
 70

561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600
 75

5 601 GGAAACTACACCGACCACGCTGTTGGTACAACACTG 640
 .
 .
 10 641 GCTTGGAGCGTGTCTGGGTCTGATTCTAGAGATTGGAT 680
 .
 .
 15 681 TAGATAACAACCAGTTCAAGGAGAGAATTGACCCCTCACAGTT 720
 .
 .
 20 721 TTGGACATTGTGTCTCTTCCCAGACTATGACTCCAGAA 760
 .
 .
 25 761 CCTACCCCTATCCGTACAGTGTCCCAACTTACCAAGAGAAAT 800
 .
 .
 30 801 CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC 840
 .
 .
 35 841 CGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAGCC 880
 .
 .
 40 881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920
 .
 .
 45 921 CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAAG 960
 .
 .
 50 961 ATCATGGCCTCTCCAGTTGGATTCAAGCGGGCCCGAGTTA 1000
 .
 .
 55 1001 CCTTTCTCTATGAACTATGGAAACACGCCGCTCCACA 1040
 .
 .
 60 1041 ACAACGTATCGTTGCTCAACTAGGTCAAGGGTGTCTACAGA 1080
 .
 .
 65 1081 ACCTTGTCTTCCACCTTGTACAGAAGACCCCTCAATATCG 1120
 .
 .
 70 1121 GTATCAACAACCAGCAACTTCCGTTCTGACGGAACAGA 1160
 .
 .
 75 1161 GTTCGCCTATGGAACCTCTTAACCTGCCATCCGCTGTT 1200

5	1201	TACAGAAAGAGCGGAACC GTTGATTCCCTGGACGAAATCC	1240
10	1241	CACCA CAGAACAAACAATGTGCCACCCAGGCAAGGATTCTC	1280
15	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTC	1320
20	1321	AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT	1360
25	1361	CTTGGATA CACCCGTAGTGCTGAGTTCAACAAACATCATCGC	1400
30	1401	ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
35	1441	TTTCTCTTCAACGGTTCTGTCAATTCAAGGACCAGGATTCA	1480
40	1481	CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA	1520
45	1521	CATTAGAAATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
50	1561	CCATCCACATCTACCA GATATAGAGTTCTGTGAGGTATG	1600
55	1601	CTTCTGTGACCCCTATTCA CCTCAACGTTAATTGGGGTAA	1640
60	1641	TTCA TCCATCTCTCCAATACAGTTCCAGCTACAGCTACC	1680
65	1681	TCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTTTG	1720
70	1721	AAAGTGCCAATGCTTTACATCTTCACTCGGTAAACATCGT	1760
75	1761	GGGTGTTAGAAACCTTAGTGGGACTGCAGGAGTGATTATC	1800

5	1801	GACAGATTCGAGTTCATTCAGTTACTGCAACACTCGAGG	1840
	1841	CTGAATATAATCTGGAAAGAGGCCAGAAGGCAGGTGAATGC	1880
10	1881	GCTGTTACGTCTACAAACCAGCTCGGCCTCAAGACCAAT	1920
15	1921	GTGACGGATTATCATATTGATCAAGTGTCCAACTTGGTGA	1960
20	1961	CCTACCTCAGCGATGAGTTCTGTCTGGATGAAAAGCGAGA	2000
25	2001	ATTGTCCGAGAAAGTCAAACATGCGAACCGACTCAGTGAT	2040
30	2041	GAACGCAATTACTCCAAGATTCAAATTCAAAGACATTA	2080
35	2081	ATAGGCAACCAGAACGTGGTGGGGCGGAAGTACAGGGAT	2120
40	2121	TACCATCCAGGGAGGTGACGACGTGTTCAAGGAGAACTAC	2160
45	2161	GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT	2200
50	2201	ACCTCTACCAGAAGATCGACGAGTCCAAGTTGAAAGCCTT	2240
	2241	TACCCGTTATCAATTAAAGAGGGTATATCGAAGATAGTCAT	2280
	2281	GACCTCGAGATCTACCTCATCCGCTACAATGCAAAACATG	2320
	2321	AAACAGTAAATGTGCCAGGTACGGGTTCTTATGGCCGCT	2360
	2361	TTCAAGCCCAGTCCAATCGGAAAGTGTGGAGGCCGAAT	2400

5	2401	CGATGCGCGCACACCTTGAATGGAATCCTGACTTAGATT	2440
	2441	GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATTCGCA	2480
10	2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
15	2521	AATGAGGACCTAGGTGTATGGGTGATCTTAAAGATTAAGA	2560
20	2561	CGCAAGATGGGCACGCAAGACTAGGAAATCTAGAGTTCT	2600
	2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
25	2641	AAAAGAGCGGAGAAAAATGGAGAGACAAACGTGAGAAGT	2680
	2681	TGGAATGGGAGACCAACATCGTCTACAAAGAGGCAAAAGA	2720
30	2721	ATCTGTAGATGCTTTATTGTAAACTCTCAATATGATCAA	2760
	2761	TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG	2800
35	2801	ATAAACGTGTTCATAGCATTGAGAAGCTTATCTGCCTGA	2840
40	2841	GCTGTCGTGATTCCGGGTGTCAATGCGGCTATTTTGAA	2880
	2881	GAATTAGAAGGGCGTATTTCACTGCATTCTCCCTCTACG	2920
45	2921	ATGCCAGAAACGTCATCAAGAACGGTGACTTCAACAATGG	2960
	2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000

50

55

5	3001	GAACAAAACAACCAACGTTGGTCCTTGTGTTCCCGGAAT	3040
10	3041	GGGAAGCAGAAGTGTACAAGAAGTTCGTGTCTGTCCGGG	3080
15	3081	TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGG	3120
20	3121	TATGGAGAAGGTTGCCTAACCAATTGAGATCGAGAACAA	3160
25	3161	ATACAGACGAACTGAAGTTAGCAACTGCGTAGAAGAGGA	3200
30	3201	AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT	3240
35	3241	GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA	3280
40	3281	ATCGAGGATATAACGAAGCTCCTCCGTACCGAGCTGATTA	3320
45	3321	TGCGTCAGTCTATGAAGAAAAATCGTATAACAGATGGACGA	3360
50	3361	AGAGAGAAATCCTTGTGAATTAAACAGAGGGTATAGGGATT	3400
55	3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
60	3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
65	3481	GAAACGGAAGGAACATTATCGTGGACAGCGTGGAAATTAC	3520
70	3521	TCCTTATGGAGGAA 3534	

F. Un gène de structure codant une protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

50

55

5	1	ATGGACAAACAACCCAAACATCAACGAATGCATTCCATA	40
10	41	ACTGCTTGAGTAACCCAGAAGTTGAAGTACITGGTGGAGA	80
15	81	ACGCATTGAAACCAGGTACACTCCCATCGACATCTCCTT	120
20	121	TCCTTGACACAGTTCTGCTCAGCGAGTTCGTGCCAGGT	160
25	161	CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGGTAT	200
30	201	CTTTGGTCCATCTCAATGGGATGCATTCTGGTGCAAATT	240
35	241	GAGCAGTTGATCAACCAGAGGATCGAAGAGAGTTGCCAGGA	280
40	281	ACCAGGCCATCTCTAGGTTGGAAGGGATTGAGCAATCTCTA	320
45	321	CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT	360
50	361	CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT	400
55	401	TCAACGACATGAACAGCGCCTTGAACCACAGCTATCCCATT	440
	441	GTTCCGCACTCCAGAACTACCAAGTTCTCTCTGTCCGTG	480
	481	TACGTTCAAGCAGCTAATCTCACCTCAGCGTGCTTCGAG	520
	521	ACGTTAGCGTGTGTTGGCAAAGGTGGGGATTGATGCTGC	560

	561	AACCATCAATAGCCGTTACAACGACCTTACTAGGGCTGATT	600
5	601	GGAAACTACACCGACCACGCTGTTGGTACAACACTG	640
	641	GCTTGGAGCGTGTCTGGGGTCCTGATTCTAGAGATTGGAT	680
10		.	.
	681	TAGATAACCAGTTCAAGGAGAGAATTGACCCCTCACAGTT	720
		.	.
15	721	TTGGACATTGTGTCCTCTTCCCAGAACTATGACTCCAGAA	760
		.	.
	761	CCTACCCCTATCCGTACAGTGTCCCCAACTTACCAAGAGAAAT	800
20	801	CTATACTAACCCAGTTCTTGAGAACCTCGACGGTAGCTTC	840
		.	.
	841	CGTGGTTCTGCCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
25		.	.
	881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
		.	.
30	921	CGATGCTCACAGAGGGAGAGTATTACTGGTCTGGACACCAG	960
		.	.
	961	ATCATGGCCTCTCCAGTTGGATTCAAGGGGCCGAGTTA	1000
		.	.
35	1001	CCTTCCTCTATGGAACATATGGAAACGCCGCTCCACA	1040
		.	.
	1041	ACAACGTATCGTTGCTCAACTAGGTCAAGGGTGTCTACAGA	1080
40		.	.
	1081	ACCTTGTCTTCCACCTTGTACAGAAGACCCCTTCAATATCG	1120
		.	.
45	1121	GTATCAACAACCAGCAACTTCCGTTCTTGACGGAACAGA	1160
		.	.
	1161	GTTGCCTATGGAACCTCTTCTAACCTGCCATCCGCTGTT	1200
50		.	.
	1201	TACAGAAAGAGCGGAACCGTTGATTCTTGGACGAAATCC	1240
		.	.
	1241	CACCAACAGAACAAACATGTGCCACCCAGGCAAGGATTCTC	1280
55		.	.
	1281	CCACAGGTTGAGGCCACGTGTCCATGTTCCGTTCCGGATTG	1320

5	1321	AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT	1360
10	1361	CTTGGATAACCGTAGTGCTGAGTTCAACAACATCATCGC	1400
15	1401	ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
20	1441	TTTCTCTTCAACGGTTCTGTCAATTTCAGGACCAGGATTCA	1480
25	1481	CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA	1520
30	1521	CATTAGAATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
35	1561	CCATCCACATCTACCAGATAAGAGTTCGTGTGAGGTATG	1600
40	1601	CTTCTGTGACCCCTATTCACCTAACGTTAATTGGGGTAA	1640
45	1641	TTCATCCATCTTCTCCAATACAGTTCCAGCTACAGCTACC	1680
50	1681	TCCTTGGATAATCTCCAATCCAGCGATTCTGGTTACTTG	1720
55	1721	AAAATGCCAATGCTTTACATCTTCACTCGTAACATCGT	1760
	1761	GGGTGTTAGAAACTTTAGTGGGACTGCAGGAGTGATTATC	1800
	1801	GACAGATTGAGTTCAATTCCAGTTACTGCAACACTCGAGG	1840
	1841	CTGAATATAATCTGAAAGAGCGCAGAAGGCCTGAATGC	1880
	1881	GCTGTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT	1920
	1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTAGTTA	1960
	1961	CGTATTTATCGGATGAATTTGTCTGGATGAAAAGCGAGA	2000
	2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
	2041	GAACGCAATTACTCCAAGATTCAAATTCAAAGACATTA	2080

5	2081	ATAGGCAACCAGAACGTGGGTGGGGCGGAAGTACAGGGAT	2120
10	2121	TACCATCCAAGGAGGGATGACGTATTTAAAGAAAATTAC	2160
15	2161	GTCACACTATCAGGTACCTTGATGAGTGCTATCCAACAT	2200
20	2201	ATTTGTATCAAAAAATCGATGAATCAAAATTAAAAGCCTT	2240
25	2241	TACCCGTTATCAATTAAAGAGGGTATATCGAAGATACTCAA	2280
30	2281	GACTTAGAAATCTATTTAATTGCTACAATGCAAAACATG	2320
35	2321	AAACAGTAAATGTGCCAGGTACGGGTTCTTATGGCCGCT	2360
40	2361	TTCAGCCCAGTCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
45	2401	CGATGCGGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
50	2441	GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATTCGCA	2480
55	2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
60	2521	AATGAGGACCTAGGTGTATGGGTGATCTTAAGATTAAGA	2560
65	2561	CGCAAGATGGGCACGCAAGACTAGGAAATCTAGAGTTCT	2600
70	2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
75	2641	AAAAGAGCGGAGAAAAATGGAGAGACAAACGTAAAAAT	2680
80	2681	TGGAATGGAAACAAATATCGTTATAAAGAGGGCAAAAGA	2720
85	2721	ATCTGTAGATGCTTTATTGTAAACTCTCAATATGATCAA	2760
90	2761	TTACAAGCGGATACGAATATTGCCATGATTGCGGGCAG	2800
95	2801	ATAAACGTGTTCATAGCATTGAGAAGCTTATCTGCCTGA	2840

2641	GCTGTCTGTGATTCCGGGTGTCAATGC GGCTATTTTGAA	2880
5	.	.
2881	GAATTAGAAGGGCGTATTTCACTGCATTCTCCCTATATG	2920
10	.	.
2921	ATGCGAGAAAATGTCATTAAAAATGGTGATTTAATAATGG	2960
15	.	.
2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000
20	.	.
3001	GAACAAAACAACCAACGTTCGGTCCITGTTGTTCCGGAAAT	3040
25	.	.
3041	GGGAAGCAGAAGTGTACAGAACAGTTCGTGTCTGTCCGGG	3080
3081	TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGGA	3120
3121	TATGGAGAAGGTTGCGTAACCATTATGAGATCGAGAACAA	3160
35	.	.
3161	ATACAGACGAACTGAAGTTAGCAACTGCGTAGAAGAGGA	3200
30	.	.
3201	AATCTATCCAATAAACACCGTAACGTGAAATGATTATACT	3240
3241	GTAAATCAAGAAGAATAACGGAGGTGCGTACACTCTCGTA	3280
35	.	.
3281	ATCGAGGATATAACGAAGCTCCTCCGTACCAAGCTGATTA	3320
40	.	.
3321	TGCGTCAGTCTATGAAGAAAATCGTATAACAGATGGACGA	3360
45	.	.
3361	AGAGAGAATCCTTGTGAATTAAACAGAGGGTATAGGGATT	3400
45	.	.
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
50	.	.
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
55	.	.
3481	GAAACGGAAAGGAACATTATCGTGGACAGCGTGGAAATTAC	3520
3521	TCCTTATGGAGGAA 3534.	

G. Un gène de structure codant une protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

5	1 ATGGACAACAACCAAACATCAACGAATGCATTCCATACA	40
10	41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA	80
15	81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG	120
20	121 TCCCTGACACAGTTCTGCTCAGCGAGTTCGTGCCAGGTG	160
25	161 CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGTAT	200
30	201 CTTTGGTCCATCTCAATGGGATGCATTCTGGTGCAAATT	240
35	241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTGCCAGGA	280
40	281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA	320
45	321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT	360
50	361 CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT	400
55	401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT	440
	441 GTTCGCAGTCCAGAACTACCAAGTTCTCTTTGTCCGTG	480
	481 TACGTTCAAGCAGCTAATCTCACCTCAGCGTGCTTCGAG	520
	521 ACGTTAGCGTGTGGGCAAAGGTGGGGATTGATGCTGC	560
	561 AACCATCAATAGCGTTACAACGACCTTACTAGGCTGATT	600
	601 GGAAACTACACCGACCACGCTGTTGTTGGTACAACACTG	640
	641 GCTTGGAGCGTGTCTGGGGCCTGATTCTAGAGAGATTGGAT	680

681	TAGATACAACCAGITCAGGAGAGAATTGACCCCTCACAGTT	720
5		
721	TTGGACATTGTGTCCTCTTCCCGAACTATGACTCCAGAA	760
10		
761	CCTACCCCTATCCGTACAGTGTCCCAACTTACCAAGAGAAAT	800
15		
801	CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC	840
20		
841	CGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
25		
881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
30		
921	CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAAG	960
35		
961	ATCATGGCCTCTCCAGTTGGATTCAAGCGGGCCCGAGTTA	1000
40		
1001	CCTTCCTCTATGGAACATATGGAAACGCCCCCTCCACAA	1040
45		
1041	ACAACGTATCGTTGCTCAACTAGGTCAAGGGTGTCTACAGA	1080
50		
1081	ACCTTGTCTTCCACCTTGTACAGAAGACCCCTTCAATATCG	1120
55		
1121	GTATCAAACAACCAGCAACTTCCGTTCTGACGGAACAGA	1160
60		
1161	GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT	1200
65		
1201	TACAGAAAAGCGGAACCGTTGATTCTTGGACGAAATCC	1240
70		
1241	CACCACAGAACAAACATGTGCCACCCAGGCAAGGATTCTC	1280
75		
1281	CCACAGGTTGAGGCCACGTGTCCATGTTCCGTTCCGGATT	1320
80		
1321	AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT	1360
85		
1361	CTTGGATACTCCGTAGTGCTGAGTTCAACAAACATCGC	1400
90		
1401	ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440

	1441	TTTCTCTCAACGGTTCTGTCATTCAAGGACCAGGATTCA	1480
5	1481	CTGGTGGAGACCTCGTTAGACTAACAGCAGTGGAAATAA	1520
10	1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
	1561	CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG	1600
15	1601	CTTCTGTGACCCCTATTCACCTAACGTTAATTGGGGTAA	1640
	1641	TTCATCCATCTCTCCAATACAGTTCCAGCTACAGCTACC	1680
20	1681	TCCTTGGATAATCTCCAATCCAGCGATTCTGGTTACTTTG	1720
	1721	AAAGTGCCAAATGCTTTACATCTTCACTCGGTAAACATCGT	1760
25	1761	GGGTGTTAGAAACTTTAGTGGGACTGCAGGAGTGATTATC	1800
	1801	GACAGATTGAGTTCATCCAGTTACTGCCAACACTCGAGG	1840
30	1841	CTGAGTACAACCTTGAGAGAGCCCAGAAGGCTGTGAACGC	1880
	1881	CCTCTTTACCTCCACCAATCAGCTGGCTTGAAAACTAAC	1920
35	1921	GTTACTGACTATCACATTGACCAAGTGTCCAACTTGGTCA	1960
	1961	CCTACCTTAGCGATGAGTTCTGCCTCGACGAGAACCGTGA	2000
40	2001	ACTCTCCGAGAAAGTTAACACGCCAACGCGTCTCAGCGAC	2040
	2041	GAGAGGAATCTCTTGCAAGACTCCAACTTCAAAGACATCA	2080
45	2081	ACAGGCAGCCAGAACGTGGTTGGGTGGAAGCACCAGGAT	2120
	2121	CACCATCCAAGGAGGCCAGGATGTGTTCAAGGAGAACTAC	2160
50	2161	GTCACCCCTCTCCGGAACTTTCGACGAGTGCTACCCCTACCT	2200

	2201	ACTTGTACCAAGAAGATCGATGAGTCATAACTCAAAGCCTT	2240
5	2241	CACCAAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAA	2280
10	2281	GACCTTGAAATCTACTCGATCAGGTACAATGCCAAGCACG	2320
	2321	AGACCGTGAATGTCCCAGGTACTGGTCCCTCTGGCCACT	2360
15	2361	TTCTGCCAACCTCCATTGGGAAGTGTGGAGAGCCTAAC	2400
	2401	AGATGCGCTCCACACCTTGAGTGGAACTCTGACTTGGACT	2440
20	2441	GCTCCTGCAGGGATGGCGAGAAGTGTGCCACCATCTCA	2480
25	2481	TCACTTCTCCTTGGACATCGATGTGGGATGTACTGACCTG	2520
	2521	AATGAGGACCTCGGAGTCTGGTCATCTCAAGATCAAGA	2560
30	2561	CCCAAGACGGACACGCAAGACTTGGCAACCTTGAGTTCT	2600
	2601	CGAAGAGAAACCATTGGTCGGTGAAGCTCTCGCTCGTGTG	2640
35	2641	AAGAGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAAC	2680
	2681	TCGAATGGAAACTAACATCGTTACAAGGAGGCCAAAGA	2720
40	2721	GTCCCGGGATGCTTGTTCGTGAACCTCCAATATGATCAG	2760
	2761	TTGCAAGCCGACACCAACATGCCATGATCCACGCCGCAG	2800
45	2801	ACAAACGTGTGCACAGCATTGAGGCTTACTTGCCTGA	2840
	2841	GTTGTCCGTGATCCCTGGTGTGAACGCTGCCATCTTCGAG	2880
50	2881	GAACCTGAGGGACGTATCTTACCGCATTCTCCTTGTACG	2920
55	2921	ATGCCAGAAACGTCAAGAACGGTGACTTCAACAATGG	2960

5	2961	CCTCAGCTGCTGGAATGTGAAAGGTATGTGGACGTGGAG	3000
	3001	GAACAGAACAAATCAGCGTCCGTCCCTGGTTGTGCCTGAGT	3040
10	3041	GGGAAGCTGAAGTGTCCCCAAGAGGTTAGAGTCTGTCCAGG	3080
	3081	TAGAGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGA	3120
15	3121	TACGGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAA	3160
	3161	ACACCGACGAGCTTAAGTTCTCCAACGTGCGTCGAGGAAGA	3200
20	3201	AATCTATCCAAACAACACCGTTACTTGCAACGACTACACT	3240
	3241	GTGAATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTA	3280
25	3281	ACAGAGGTTACAACGAAGCTCCTTCCGTTCTGCTGACTA	3320
	3321	TGCCTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGA	3360
30	3361	CGTGAGAACCTTGCGAGTTAACAGAGGTTACAGGGACT	3400
	3401	ACACACCACTTCCAGTTGGCTATGTTACCAAGGGAGCTTGA	3440
40	3441	GTACTTTCTGAGACCGACAAAGTGTGGATCGAGATCGGT	3480
	3481	GAAACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTC	3520
45	3521	TCCTGATGGAGGAA 3534.	

50 H. Un gène de structure qui code une protéine insecticide de *B.t.t.* Comportant La séquence :

1	ATGACTGCAGACAACAACACCGAAGCCCTCGACAGTTCTA	40
5		
41	CCACTAAGGATGTTATCCAGAAGGGTATCTCCGTTGTGGG	80
10		
81	AGACCTCTGGGCGTGGTGGATTTCCTTCGGTGGAGCC	120
15		
121	CTCGTGAGCTCTATAACAAACTTCTCAACACCATTGGC	160
20		
161	CAAGCGAGGACCCTGGAAAGCATTCATGGAGCAAGTTGA	200
25		
201	AGCTCTTATGGATCAGAAGATTGCAGATTATGCCAAGAAC	240
281	AAGGCTTGGCAGAACTCCAGGGCCTTCAGAACAAATGTGG	280
30		
321	TGTTAGCTCCAGAAATCCTCACAGCCAAGGTAGGATCAGA	360
35		
361	GAGTTGTTCTCTCAAGCCGAATCCCACCTTCAGAAATTCCA	400
40		
401	TGCCTAGCTTGCTATCTCCGGTTACGAGGTTCTTTCTC	440
441	CACTACCTATGCTCAAGCTGCCAACACCCACTTGTTCTC	480
45		
481	CTTAAGGACGCTAAATCTATGGAGAAGAGTGGGGATACG	520
521	AGAAAGAGGACATTGCTGAGTTCTACAAGCGTCAACTTAA	560
55		
561	GCTCACCCAAAGAGTACACTGACCATTGCGTGAAATGGTAT	600
601	AACGTTGGTCTCGATAAGCTCAGAGGCTTCTACGAGT	640
641	CTTGGGTGAACCTCAACAGATAACAGGAGAGAGATGACCTT	680

5	681	GACTGTGCTCGATCTTATCGCACTCTTCCTTGACGAT	720
10	721	GTGAGACTCTACCCAAAGGAAGTGAAAAGTGAGCTTACCA	760
15	761	GAGACGTGCTCACTGACCCATTGTCGGAGTCACAAACCT	800
20	801	TAGGGTTATGGAACCTACCTTCAGCAATATCGAAAACCTAC	840
25	841	ATTAGGAAACCACATCTCTCGACTATCTCACAGAACATTG	880
30	881	AATTCCACACAGGTTCAACCAGGATACTATGGTAACGAA	920
35	921	CTCCTTCAACTATTGGTCGGTAACTATGTTCCACCAGA	960
40	961	CCAAGCATTGGATCTAATGACATCATCACATCTCCCTTCT	1000
45	1001	ATGGTAACAAGTCCAGTGAACCTGTGCAGAACCTTGAGTT	1040
50	1041	CAACGGCGAGAAAGTCTATAGAGCCGTGCAAACACCAAT	1080
55	1081	CTCGCTGTGTGGCCATCCGAGTTACTCAGGGGTACAAA	1120
	1121	AGGTGGAGTTAGTCAGTATAACGATCAGACCGATGAGGC	1160
	1161	CAGCACCCAGACTTACGACTCCAAACGTAACGTTGGCGCA	1200
	1201	GTCTCTGGGATTCTATCGACCAATTGCCTCCAGAAACCA	1240
	1241	CAGACGAACCAATTGGAGAAGGGCTACAGCCACCAACTAA	1280
	1281	CTATGTGATGTGCTTCTTGATGCAAGGTTCCAGAGGGACC	1320
	1321	ATTCCAGTGTGACCTGGACACACAAGTCCGTGGACTTCT	1360
	1361	TCAACATGATCGATAGCAAGAAGATCACTCAACTTCCCTT	1400
	1401	GGTGAAAGCCTACAAGCTGCAATCTGGTGCTTCCGGTTGTC	1440

5	1441	GCAGGTCCCAGATTCACTGGAGGTGACATCATCCAGTGCA	1480
	1481	CAGAGAACGGCAGCGCAGCTACTATCTACGTGACACCTGA	1520
10	1521	TGTGTCTTACTCTCAGAAAGTACAGGGCACGTATTCAATTAC	1560
15	1561	GCATCTACCAGCCAGATCACCTTCACACTCAGCTTGGATG	1600
20	1601	GAGCACCCCTCAACCAGTATTACTTTGACAAGACCATCAA	1640
25	1641	CAAAGGTGACACTCTCACATAACAATAGCTCAACTTGGCA	1680
	1681	AGTTTCAGGCACACCATTGAACTCTCAGGCAACAATCTTC	1720
	1721	AGATCGGGGTCAACCGGTCTCAGCGCCGGAGACAAAGTCTA	1760
30	1761	CATCGACAAGATTGAGTTCATCCCAGTGAAC	1791

35 I. Un gène de structure qui code une protéine insecticide de *B.t.* entomocidus comportant la séquence :

40

45

50

55

5	1	ATGGAGGAGAACAAACCAAAACCAATGCATTCCATACAAC	40
10	41	GCTTGAGTAACCCAGAAGAGGTATTGCTTGATGGAGAACG	80
15	81	CAATTCAACCGGTAACTCTTCCATCGACATCTCCTGTCC	120
20	121	TTGGTCCAGTTCTGGTCAGCAACTTCGTGCCAGGTGGTG	160
25	161	GGTTCCCTTGTCCGGACTAATTGACTTCGTCTGGGTATCGT	200
30	201	TGGTCCATCTCAATGGGATGCATTCTGGTCAAATTGAG	240
35	241	CAGTTGATCAACGAGAGGATCGCTGAGTTGCCAGGAACG	280
40	281	CTGCCATCGCTAACTTGAAGGATTGGCAATAACTCAA	320
45	321	CATCTATGTGGAGGCCTTCAAAGAGTGGGAAGAGGACCC	360
50	361	AACAACCCAGAGACCGCACTAGGGTATCGACAGATTCA	400
55	401	GAATCTTGGACGGCCTCTGGAGAGAGATATCCCATCCTT	440
	441	CAGAAATCTCTGGCTTCGAAGTTCCCTCTCTGTCCGTGTAC	480

481	GCTCAAGCAGCTAATCTTACACCTGGCTATCCTTCGAGACA	520
521	GTGTCATCTTGGGGAAAGGTGGGGATTGACCACTATCAA	560
561	CGTCAATGAGAATTACAACAGACTTATCAGGCACATTGAC	600
601	GAGTACGCCGACCACTGTGCTAACACCTACAAACCGTGGCT	640
641	TGAACAATCTCCCTAACAGTCTACTTATCAAGATTGGATTAC	680
681	CTACAACAGGTTGAGGAGAGACTTGACCCCTCACAGTTTG	720
721	GACATTGCAGCTTCTTCCCAGACTATGACAACAGGAGAT	760
761	ACCCATCCAACCAGTGGGTCAACTTACCAAGAGAAGTCTA	800
801	TACTGACCCACTTATCAACTTCAACCCCTCAGTTGCAAAGT	840
841	GTCGCCCAACTTCCCACATTCAACGTCACTGGAGTCCAGCC	880
881	GTATCAGGAACCCACACTTGTGACATCTTGAACAACCT	920
921	TACTATCTTACCGATTGGTTCAAGCGTTGGCGTAACCTC	960
961	TATTGGGGTGGACACAGGGTCATCTCCTCTTTATTGGAG	1000
1001	GTGGGAACATTACCTCTCCTATCTATGGACGTGAGGCAAA	1040
1041	CCAGGAGCCACCACGTAGTTCACCTTCAACGGTCCAGTC	1080
1081	TTCAGAACCTTGTCTAACCTTACCTTGAGATTGCTCCAGC	1120
1121	AACCTTGGCCAGCTCCACCTTCAACCTTAGAGGTGTTGA	1160
1161	GGGCGTTGAGTTCTCTACTCCTACCAACTCCTCACTTAC	1200
1201	AGAGGTAGAGGAACCGTTGATTCTTGACCGAACTCCCAC	1240

5	1241	CAGAGGACAATAGCGTGCCACCCAGGGAAGGCTACTCCCCA	1280
10	1281	CAGGGTGTGCCACGCAACCTTCGTGCAGCGTCCGGAACT	1320
15	1321	CCATTCCCTCACTACAGGAGTTGTGTTCTCATGGACTGATC	1360
20	1361	GTAGTGCTACTCTCACTAACCAATTGATCCCGAGAGGAT	1400
25	1401	CAATCAAATCCCATTGGTCAAGGGTTCCGTGTGTGGGGA	1440
30	1441	GGAACCTCTGTACATCACAGGACCAGGCTTCACAGGAGGTG	1480
35	1481	ATATTCTTAGAAGAAAACACTTTGGCGACTTTGTGAGCCT	1520
40	1521	CCAAGTTAACATCAACTCTCCAATTACTCAAAGATAATCGT	1560
45	1561	CTCAGGTTTCGTTACGCATCTTCCCGTGACGCTAGAGTCA	1600
50	1601	TCGTGCTCACCGGAGCAGCTTCTACCGGTGTCGGTGGACCA	1640
55	1641	AGTCTCCGTGAACATGCCACTCCAGAAGACTATGGAGATC	1680
	1681	GGCGAGAACTTGACATCCAGGACCTTCAGATAACACCGACT	1720
	1721	TCTCTAACCCCTTCAGTTCCGTGCCAACCTGACATCAT	1760
	1761	TGGCATTAGCGAACACCTCTCTTTGGAGCTGGTAGCATC	1800
	1801	TCATCTGGCGAATTGTACATTGACAAGATTGAGATCATTC	1840
	1841	TTGCCGACGCTACCTTCGAGGCTGAGTCTGACCTTGAGAG	1880
	1881	AGCCCAGAAGGCTGTGAACGCCCTCTTACCTCCTCTAAT	1920
	1921	CAGATTGGCTTAAAAACTGACGTTACTGACTATCACATTG	1960
	1961	ACCAAGTGTCCAACCTGGTCGACTGCCTTAGCGATGAGTT	2000

	2001	CTGGCTCGACGAGAAGCGTGAACCTCTCGAGAAAGTTAAA	2040
5	2041	CACGCCAACGCGTCTCAGCGACGAGAGGAATCTCTTGCAG	2080
	2081	ACCCCAACTTCAGAGGCATCACAGGCAGCCAGACCGTGG	2120
10	2121	TTGGAGAGGAAGCACCGACATCACCATCCAAGGAGGCGAC	2160
	2161	GATGTGTTCAAGGAGAACTACGTCACCCCTCCCAGGAAC	2200
15	2201	TGGACGAGTGCTACCTAACCTACTTGTACCAGAAAGATCGA	2240
20	2241	TGAGTCAAACCTAAAGCCTACACCAGGTATGAACCTAGA	2280
	2281	GGCTACATCGAAGACAGCCAAGACCTTGAAATCTACCTCA	2320
25	2321	TCAGGTACAATGCCAACGACAGAGATCGTGAATGTCCCAGG	2360
	2361	TACTGGTTCCCTCTGGCCACTTTCTGCCAAATGCCATT	2400
30	2401	GGGAAGTGTGGAGAGCCTAACAGATGCGCTCACACCTTG	2440
	2441	AGTGGAAATCCTGACTTGGACTGCTCCTGCAGGGATGGCGA	2480
35	2481	GAAGTGTGCCACCATTCTCATCACTTCACCTTGGACATC	2520
40	2521	GATGTGGATGTACTGACTGAAATGAGGACCTCGGAGTCT	2560
	2561	GGGTCACTTCAAGATCAAGACCCAAAGACGGACACGCAAG	2600
45	2601	ACTTGGCAACCTTGAGTTCTCGAAGAGAGAAACCATTGCTC	2640
	2641	GGTGAAGCTCTCGCTCGTGTGAAGAGAGCAGAGAAGAGT	2680
50	2681	GGAGGGACAAACGTGAGAAACTCCAACTCGAGACTAACAT	2720
	2721	CGTTTACAAGGAGGCCAAAGAGTCCGTGGATGCTTGTTC	2760

	2761	GTGAACCTCCAATATGATAAGGTGCAAGTGGACACCPACA	2800
5	2801	TCGCCATGATCCACGCTGCAGACAAACGTGTGCACAGGAT	2840
	2841	TCTGTGAGGCTTACTTGCCCTGAGTTGTCCGTATCCCTGGT	2880
10	2881	GTGAACGCTGCCATCTCGAGGAACCTTGAGGGACGTATCT	2920
	2921	TTACCGCATACTCCTTGTACGATGCCAGAAACGTCAATCAA	2960
	2961	GAACGGTGAATTCAACAAATGGCCTCTTGTGCTGGAATGTG	3000
20	3001	AAAGGTCAATGTGGACGTGGAGGAACAGAACAAATCACCGTT	3040
	3041	CCGTCTGGTATCCCTGAGTGGGAAGCTGAAGTGTCCCA	3080
25	3081	AGAGGTTAGAGTCTGTCCAGGTAGAGGCTACATTCTCCGT	3120
	3121	GTGACCGCTTACAAGGAGGGATACTGGTGGGGTTGCGTGA	3160
30	3161	CCATCCACGAGATCGAGGACAACACCCGACGAGCTTAAGTT	3200
	3201	CTCCAAC TCGGT CGAGGAAGAAGTCTATCCCACAAACACC	3240
	3241	GTTACTTGCAACAACTACACTGGGACCCAGGAAGAGTACG	3280
40	3281	AAGGTACCTACACTAGCCGTAACCAAGGTTACGACGAAGC	3320
	3321	TTACGGAAACAATCCCTCCGTTCTGCTGACTATGCCTCC	3360
45	3361	GTGTACGAGGAGAAATCCTACACAGATGGCAGACGTGAGA	3400
	3401	ACCCCTTGCAGTCCAACACAGAGGTTACGGTACTACACACC	3440
	3441	ACTTCCAGCAGGCTATGTTACCAAGGACCTTGAGTACTTT	3480
55	3481	CCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAAACCG	3520

3521 AGGGTACCTTCATCGTGGACAGCGTGGAGCTTCCTTGAT 3560

5

3561 GGAGGAA 3567.

10 J. Un gène de structure qui code - une protéine insecticide P2 comportant la séquence :

15

20

25

30

35

40

45

50

55

5	1	ATGGACAAACAACGTCTTGAACCTCTGGTAGAACGACCCATCT	40
10	41	GCGACGCATAACAACGTCTGGCTCACGATCCATTGAGCTT	80
15	81	CGAACACAAGAGCCTCGACACTATTCAAGAAGGAGTGGATG	120
20	121	GAATGGAAACGTACTGACCACTCTCTACGTGCGACCTG	160
25	161	TGGTTGGAACAGTGTCCAGCTTCCTCTCAAGAAGGTCGG	200
30	201	CTCTCTCATCGAAAACGTATCTGTCCGAACACTCTGGGT	240
35	241	ATCATCTTCCATCTGGTCCACTAATCTCATGCAAGACA	280
40	281	TCTTGAGGGAGACCGAACAGTTCTCAACCAGCGTCTCAA	320
45	321	CACTGATACCTTGGCTAGAGTCAACGCTGAGTTGATCGGT	360
50	361	CTCCAAGCAAACATTGAGTTCAACCAGCAAGTGGACA	400
55	401	ACTTCTTGAATCAAACACTCAGAACATCCTGTGCCCTTTCCAT	440
	441	CACTTCTTCCGTGAACACTATGCAGCAACTCTTCCCTCAAC	480
	481	AGATTGCCCTCAGTTCAAGGCTACCGAGTTGCTCC	520
	521	TTCTTCCACTCTTGCTCAGGCTGCCAACATGCACTTGTC	560
	561	CTTCATACGTGACGTGATCCTCAACGCTGACGAATGGGA	600

5	601	ATCTCTGCAGCCACTCTAGGACATAACAGAGACTACTTGA	640
10	641	GGAACTAACACTCGTATTACTCCAACATTGGCATCAACAC	680
15	681	TTATCAGACTGCCTTCGTGGACTCAAATACTAGGCTTCAC	720
20	721	GACATGCTTGAGTTCAAGGACCTACATGTTCTTAACGTGT	760
25	761	TTGAGTACGTCAGCATTGGAGTCTCTCAAGTACCAAGAG	800
30	801	CTTGATGGTGTCCCTCTGGAGCCAATCTCTACGCCCTGGC	840
35	841	AGTGGACCACAGCAAACCTCAGAGCTTCACAGCTCAGAACT	880
40	881	GGCCATTCTGTATAGCTTGTCCAAGTCAACTCCAACCA	920
45	921	CATTCTCAGTGGTATCTCTGGGACCAAGACTCTCCATAACC	960
50	961	TTTCCCAACATTGGTGGACTTCCAGGCTCCACTACAACCC	1000
55	1001	ATAGCCTTAACCTCTGCCAGAGTGAACACTACAGTGGAGGTG	1040
60	1041	CAGCTCTGGATTGATTGGTGCAACTAACCTGAACCACAAAC	1080
65	1081	TTCAATTGCTCCACCGTCTGCCACCTCTGAGCACACCGT	1120
70	1121	TTGTGAGGTGCTGGCTTGACAGCGGTACTGATCGCGAAGG	1160
75	1161	AGTTGCTACCTCTACAAACTGGCAAACCGAGTCCTTCCAA	1200
80	1201	ACCACTCTTAGCCTCGGTGTGGAGCTTCTGCACGTG	1240
85	1241	GGAATTCAAACACTTTCCAGACTACTTCATTAGGAACAT	1280
90	1281	CTCTGGTGTCTCGTCATCAGGAATGAAGACCTCACC	1320
95	1321	CGTCCACTTCATTACAAACAGATTAGGAACATCGAGTCTC	1360

5 1361 CATCCGGTACTCCAGGAGGTGCAAGAGAGCTTACCTCGTGTC 1400

10 1401 TGTCCATAACAGGAAGAACAAACATCTACGCTGCCAACGAG 1440

15 1441 AATGGCACCATGATTCACCTTGCACCAGAAGATTACACTG 1480

20 1481 GATTCAACCATTCTCCAATCCATGCTACCCAAGTGAACAA 1520

25 1521 TCAGACACGCACCTTCATCTCCGAAAAGTTCGGAAATCAA 1560

30 1561 GGTGACTCCTTGAGGTTCGAGCAATCCAACACTACCGCTA 1600

35 1601 GGTACACTTGAGAGGCAATGGAAACAGCTACAACCTTTA 1640

40 1641 CTTGAGAGTTAGCTCCATTGGTAACCTCCACCATCCGTGTT 1680

45 1681 ACCATCAACGGACGTGTTACACAGTCTCTAAATGTGAACA 1720

50 1721 CTACAACGAACAATGATGGCGTTAACGACAACGGAGCCAG 1760

55 1761 ATTCAAGCGACATCAACATGGCAACATCGTGGCCTCTGAC 1800

60 1801 AACACTAACGTTACTTGAGACATCAATGTGACCCCTCAATT 1840

65 1841 CTGGAACTCCATTGATCTCATGAACATCATGTTGTGCC 1880

70 1881 AACTAACCTCCCTCCATTGTAC 1902 ou

50 K. Une séquence de gène de structure codant une protéine de fusion comprenant les acides aminés 610 N-terminaux de *B.t.k.* HD-1 et les acides aminés 567 C-terminaux de *B.t.k.* HD-73, ledit gène comportant la séquence :

1	ATGGACAACAAACCCAAACATCAACGAATGCATTCATACA	40
5		
41	ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA	80
10		
81	ACCGATTGAAACCGGTTACACTCCCCATCGACATCTCCTTG	120
15		
121	TCCTTGACACAGTTCTGCTCAGCGAGTTCGTGCCAGGTG	160
20		
161	CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGTAT	200
25		
201	CTTTGGTCCATCTCAATGGATGCATTCTGGTGCAAATT	240
30		
241	GAGCAGTTGATCAACCAGAGGATCGAAGAGTTGCCAGGA	280
35		
281	ACCAGGCCATCTCTAGGTTGGAAGGGATTGAGCAATCTCTA	320
40		
321	CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT	360
45		
361	CCTACTAACCCAGCTCTCCGCAGGAAATGCGTATTCAAT	400
50		
401	TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT	440

441 GTTCGGCAGTCCAGAACTACCAAGTTCTCTCTGTCCGTG 480
5
481 TACGTTCAAGCAGCTAACATCTCACCTCAGCGTGCTCGAG 520
10
521 ACGTTAGCGTGTGGGCAAAGGTGGGGATTCGATGCTGC 560
15
561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600
20
601 GGAAACTACACCGACCACGCTGTTGTTGGTACAACACTG 640
641 GCTTGGAGCGTGTCTGGGGTCTGATTCTAGAGATTGGAT 680
25
681 TAGATACAACCAGTTCAAGGAGAGAATTGACCCCTCACAGTT 720
721 TTGGACATTGTGTCTCTTCCCCGAACATGACTCCAGAA 760
761 CCTACCCCTATCCGTACAGTGTCCCCAACTTACCAAGAGAAAT 800
30

35

40

45

50

55

5	801	CTATACTAACCCAGTTCTTGAGAACCTCGACGGTAGCCTC	840
10	841	CGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
15	881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
20	921	CGATGTCACAGAGGAGAGTATTACTGGTCTGGACACCAAG	960
25	961	ATCATGGCCTCTCCAGTTGGATTCAAGCGGGCCCGAGTTA	1000
30	1001	CCTTTCCCTCTATGAACTATGGAAACGCCGCTCCACA	1040
35	1041	ACAACGTATCGTTGCTCAACTAGGTCAAGGGTGTCTACAGA	1080
40	1081	ACCTTGTCTTCCACCTTGTACAGAAGACCCCTCAATATCG	1120
45	1121	GTATCAACACCAGCAACTTCCGTTCTGACGGAACAGA	1160
50	1161	GTTCGCCTATGAAACCTCTTCTAACTTGCCATCCGCTGTT	1200
55	1201	TACAGAAAGAGCGGAACCGTTGATTCCCTGGACGAAATCC	1240
	1241	CACCACAGAACACAATGTGCCACCCAGGCAAGGATTCTC	1280
	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTC	1320
	1321	AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT	1360
	1361	CATGGATTCATCGTAGTGCTGAGTTCAACAAATATCATCC	1400
	1401	TTCCTCTCAAATCACCCAAATCCCATTGACCAAGTCTACT	1440
	1441	AACCTTGGATCTGGAACCTCTGTGCGTGAAAGGACCAGGCT	1480
	1481	TCACAGGAGGTGATATTCTTAGAAGAACCTCTCCTGGCCA	1520

5	1521	GATTAGCACCCCTCAGAGTTAACATCACTGCACCACCTTCT	1560
10	1561	CAAAGATATCGTGTCAAGGATTCGTTACGCATCTACCACTA	1600
15	1601	ACTTGCAATTCCACACCTCCATCGACGGAAGGCCTATCAA	1640
20	1641	TCAGGGTAACCTCTCCGCAACCATGTCAAGCGGCAGCAAC	1680
25	1681	TTGCAATCCGGCAGCTTCAGAACCGTCGGTTTCACTACTC	1720
30	1721	CTTTCAACTTCTCTAACGGATCAAGCGTTTACCCCTTAG	1760
35	1761	CGCTCATGIGTTCAATTCTGGCAATGAAGTGTACATTGAC	1800
40	1801	CGTATTGAGTTGTGCCTGCCGAAGTTACCCCTCGAGGCTG	1840
45	1841	AGTACAACCTTGAGAGAGGCCAGAACGGCTGTGAACGCCCT	1880
50	1881	CTTTACCTCCACCAATCAGCTTGGCTTGAAGAACTAACGTT	1920
55	1921	ACTGACTATCACATTGACCAAGTGTCCAACTTGGTCACCT	1960
	1961	ACCTTAGCGATGAGTTCTGCCTCGACGAGAACGGTGAACCT	2000
	2001	CTCCGAGAAAGTTAACACACGCCAACCGTCTCAGCGACGAG	2040
	2041	AGGAATCTCTTGCAAGACTCCAACCTCAAAGACATCAACA	2080
	2081	GGCAGCCAGAACGTGGTGGGTGGAAGCACCGGGATCAC	2120
	2121	CATCCAAGGAGGCACGATGTGTTCAAGGAGAACTACGTC	2160
	2161	ACCCCTCTCCGGAACCTTCGACGAGTGTGCTACCCCTACCTACT	2200
	2201	TGTACCAGAACGATCGATGAGTCCAACCTCAAAGCCTTCAC	2240
	2241	CAGGTATCAACTTAGAGGGTACATCGAAGACAGCCAAGAC	2280

	2281	CTTGAATCTACTCGATCAGGTACAATGCCAAGCACGAGA	2320
5	2321	CCGTGAATGTCCCAGGTACTGGTCCCTCTGGCCACTTTC	2360
10	2361	TGCCCAATCTCCCATTGGGAAGTGTGGAGAGCCTAACAGA	2400
	2401	TGCCTCCACACCTTGAGTGGAACTCCTGACTTGGACTGCT	2440
15	2441	CCTGCAGGGATGGCGAGAAGTGTGCCACCATTCTCATCA	2480
	2481	CTTCCTTGGACATCGATGTGGATGTACTGACCTGAAT	2520
20	2521	GAGGACCTCGGAGTCTGGTCATCTCAAGATCAAGACCC	2560
	2561	AAGACGGACACGCAAGACTTGGCAACCTTGAGTTCTCGA	2600
25	2601	AGAGAAAACCATTGGTCGGTGAAGCTCTCGCTCGTGTGAAG	2640
	2641	AGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAAACTCG	2680
30	2681	AATGGGAAACTAACATCGTTACAAGGAGGCCAAAGAGTC	2720
	2721	CGTGGATGCTTGTTCGTGAACCTCCAATATGATCAGTTG	2760
35	2761	CAAGCCGACACCAACATGCCATGATCCACGCCAGACA	2800
	2801	AACGTGTGCACAGCATTGAGGCTTACTTGCCTGAGTT	2840
40	2841	GTCCGTGATCCCTGGTGTGAACGCTGCCATTTGAGGAA	2880
	2881	CTTGAGGGACGTATCTTACCGCATCTCCTTGTACGATG	2920
45	2921	CCAGAAACGTCAAGAACGGTGAATTCAACAATGGCCT	2960
	2961	CAGCTGCTGGAATGTGAAAGGTATGTGGACGTGGAGGAA	3000
50	3001	CAGAACAAATCAGCGTTCCGTCTGGTTGTGCCTGAGTGGG	3040

5	3041	AAGCTGAAGTGTCCCAGGTTAGAGTCTGTCCAGGTAG	3080
10	3081	AGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGATAC	3120
15	3121	GGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAAACA	3160
20	3161	CCGACGAGCTTAAGTTCTCCAAC TGCGTCGAGGAAGAAAT	3200
25	3201	CTATCCCACAAACACCGTTACTTGCACGACTACACTGTG	3240
30	3241	AATCAGGAAGAGTACGGAGGTGCCTACACTAGGCCATAACA	3280
35	3281	GAGGTTACAAACGAAGCTCCTCCGTTCTGCTGACTATGC	3320
40	3321	CTCCGTGTACGAGGGAGAAATCCTACACAGATGGCAGACGT	3360
45	3361	GAGAACCCTTGCGAGTTCAACAGAGGTTACAGGGACTACA	3400
50	3401	CACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGAGTA	3440
55	3441	CTTTCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAA	3480
	3481	ACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTCTCT	3520
	3521	TGATGGAGGAA 3531.	

Determination of DNA regions in genes to modify by site-directed mutagenesis for increased expression in plants

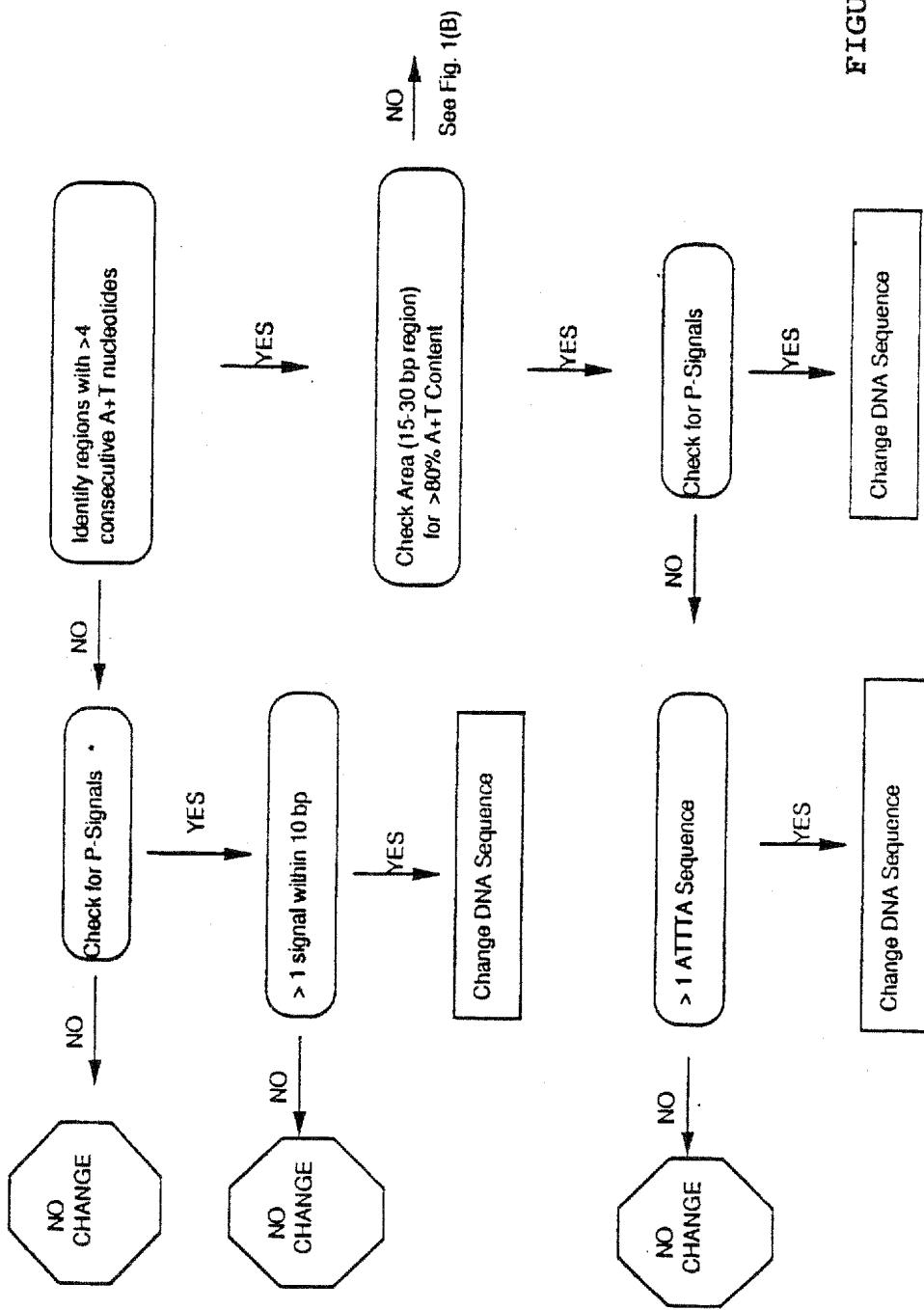
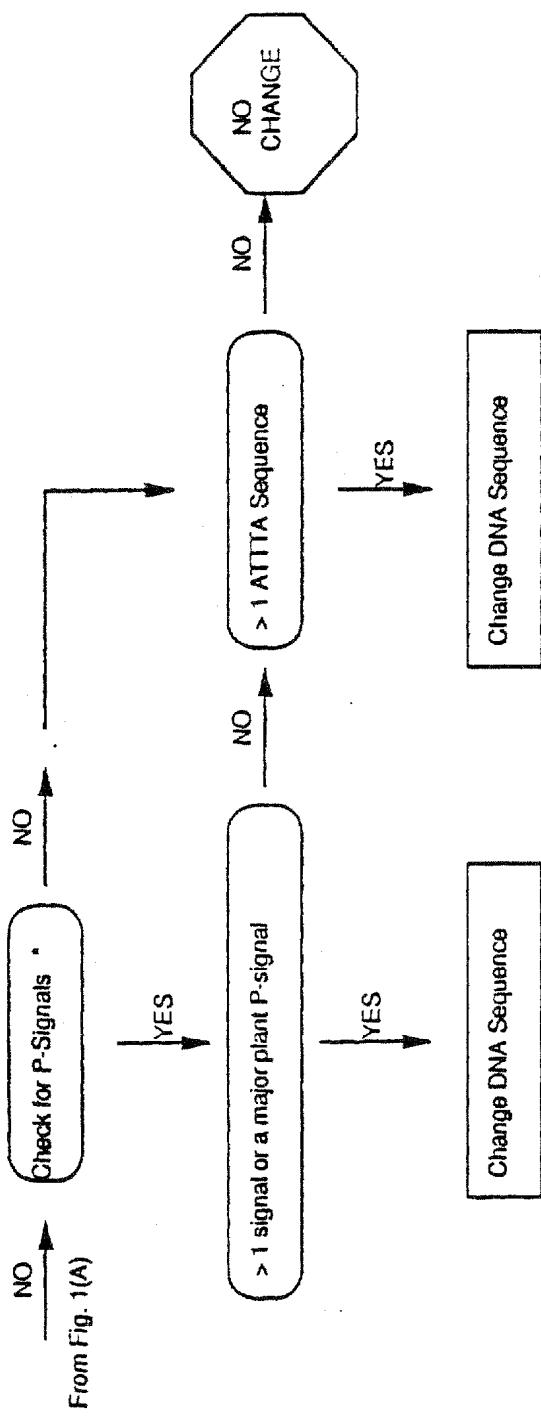


FIGURE 1 (A)

Polyadenylation Signal Sequences

Determination of DNA regions in genes to modify by
site-directed mutagenesis for increased expression in plants



* Polyadenylation Signal Sequences

FIGURE 1 (B)

1	ATGGCTATAAGAAA CTGGTTACACCCCAATCGATATTCTCT	40
41	TGTCGCTAACGCAATTCTTGTGAGTGAATTGTTCCC GG	80
81	TGCTGGATTGTGTTAGGACTAGTTGATATAATATGGGG A T C	120
121	ATTTTGGTCCCTCTCAATGGGACGCATTCTGTACAAA	160
161	TTGAACAGTTAACCAAAAGAATAGAAGAATT CC C G C G CGTAG	200
201	GAACCAAGCCATTCTAGATTAGAAGGACTAAGCAATCTT T	240
241	TATCAAATTACGCAGAACATTAGAGAGTGGGAAGCAG	280
281	ATCCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCA	320
321	ATTCAATGACATGAACAGTGCCCTTACAACCGCTATTCT	360
361	CTTTTGCAGTTCAAAATTATCAAGTTCCCTTTTATCAG CC C C	400
401	TATATGTTCAAGCTGCAAATTACATTATCAGTTTGAG G C C CC C CC C	440
441	AGATGTTCAAGTGGACAAAGGTGGGGATTGATGCC	480
481	GCGACTATCAATAGTCGTTATAATGATTAACTAGGCTTA	520
521	TTGGCAACTATACAGATCATGCTGTACGCTGGTACAATAC	560
561	GGGATTAGAGCGTGTATGGGGACCGGATTCTAGAGATTGG	600
601	ATAAGATATAATCAATTAGAAGAGAATTAAACACTA CTG C G C C G C GC T	640
641	TATTAGATATCGTTCTCTATTCCGAACTATGATAGTAG	680
681	AACGTATCCAATTGAAACAGTTCCCAATTAAACAAGAGAA	720

FIGURE 2A

721	ATTTATACAAACCCAGTATTAGAAAATTTGATGGTAGTT	760
761	TTCGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAG	800
801	TCCACATTTGATGGATATACTTAATAGTATAACCATCTAT	840
841	ACGGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATC C C C T C	880
881	AAATAATGGCTTCTCCTGTAGGGTTTCGGGCCAGAATT G C	920
921	CACTTTCCGCTATATGGAACTATGGGAAATGCAGCTCCA	960
961	CAACAACGTATTGTTGCTCAACTAGGTCAAGGGCGTGTATA	1000
1001	GAACATTATCGTCCACCTTATATAGAACGACCTTTAATAT C	1040
1041	AGGGATAAATAATCAACAACTATCTGTTCTTGACGGGACA C C C C	1080
1081	GAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTG	1120
1121	TATACAGAAAAAGCGGAACGGTAGATTGCTGGATGAAAT	1160
1161	ACCGCCACAGAATAACAAACGTGCCACCTAGGCAAGGATT	1200
1201	AGTCATCGATTAAGCCATGTTCAATGTTCGTCAGGCT	1240
1241	TTAGTAATAGTAGTGTAAAGTATAATAAGAGCTCCTATGTT	1280
1281	CTCTGGATACATCGTAGTGCTGAATTTAATAATATAATT G C C C C C	1320
1321	CCTTCATCACAAATTACACAAAATACCTTTAACAAAATCTA C C C AC C C G	1360
1361	CTAATCTTGGCTCTGGAACCTCTGTCGTTAAAGGACCAGG	1400

FIGURE 2B

1401 ATTTACAGGAGGAGATATTCTCGAAGAACCTCACCTGGC 1440
1441 CAGATTCAACCTTAAGAGTAAATATTACTGCACCATTAT 1480
1481 CACAAAGATATCGGGTAAGAATTGCTACGCTTCTACCAC 1520
1521 AAATTTACAATTCCATACATCAATTGACGGAAGACCTATT
 CC T G C 1560
1561 AATCAGGGGAATTTTCAGCAAATGAGTAGTGGAGTA 1600
1601 ATTTACAGTCCGGAAGCTTAGGACTGTAGGTTTACTAC 1640
1641 TCCGTTAACCTTCAAATGGATCAAGTGTATTTACGTTA 1680
1681 AGTGCTCATGTCTCAATTCAAGGCAATGAAGTTATATAG 1720
1721 ATCGAATTGAATTGTTCCGGCA 1743

FIGURE 2C

1	ATGGATAACAATCCGAACATCAATGAATGCATTCTTATA	40
	C C A C A C	
41	ATTGTTAACGTAACCCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAAACTGGTTACACCCCAATCGATATTCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTCTTTGAGTGAATTGTTCCCGGTG	160
	CT G A G G C C G C G A	
161	CTGGATTGTTAGGACTAGTTGATATAATATGGGAAT	200
	G C T C C C C T	
201	TTTGGTCCCTCTCAATGGGACGCATTCTTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAACCAACAAAGAATAGAAGAATTGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTCTAGATTAGAAGGACTAAGCAATCTTA	320
	G C G G T G C	
321	TCAAATTACGCAGAACATCTTTAGAGAGTGGGAAGCAGAT	360
	C C T G A C	
361	CCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAAT	400
	C T C C C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCTCT	440
	C C T G C A C AT	
441	TTTGCAGTTCAAAATTATCAAGTTCTCTTTATCAGTA	480
	G C C G C C C G G	
481	TATGTTCAAGCTGCAAATTACATTATCAGTTTGAGAG	520
	C A T C T C C C A G C G C T C	
521	ATGTTTCAGTGTGGACAAAGGTGGGGATTTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACCTAGGCTTATT	600
	A C C C C C C T G	
601	GGCAACTATACAGATCATGCTGTaCGCTGGTACAATACGG	640
	A C C C C T T C T	
641	GATTAGAGCGTGTATGGGGACCGGATTCTAGAGATTGGAT	680
	C G C T T	

FIGURE 3A

681	AAGATATAATCAATTAGAAGAGAATTAACACTAACTGTA	720
	T C C G C G G C C A T	
721	TTAGATATCGTTCTCTATTTCGAACATATGATAGTAGAA	760
	G C T G C C C T C C	
761	CGTATCCAATTGAAACAGTTCCCAATTAAACAAGAGAAAT	800
	C C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAATTTGATGGTAGTTT	840
	C T T C T G C C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAGTC	880
	T T C A T C C T C C C	
881	CACATTGATGGATATACTTAATAGTATAACCATCTATAAC	920
	C C T G C C T C	
921	GGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATCAA	960
	C C G C T A C G	
961	ATAATGGCTTCTCCTGTAGGGTTTCGGGCCAGAATTCA	1000
	C C A T A C A G C C G T	
1001	CTTTCCGCTATATGAACTATGGAAATGCAGCTCCACAC	1040
	C T C C C	
1041	ACAACGTATTGTTGCTCAACTAGGTCAAGGGCGTGATAGA	1080
	C T C C	
1081	ACATTATCGTCCACCTTATATAGAACGACCTTTAATATAG	1120
	C G T G C C C C	
1121	GGATAAAATAATCAACAACTATCTGTTCTGACGGGACAGA	1160
	T C C C G T C A	
1161	ATTTGCTTATGGAACCTCCTCAAATTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTGCTGGATGAAATAC	1240
	G C T C T C C	
1241	CGCCACAGAATAACAAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C C T C C	
1281	TCATCGATTAAGCCATGTTCAATGTTGTTCAAGGCTTT	1320
	C C A G G C G C C C A C	
1321	AGTAATAGTAGTGTAAAGTATAATAAGAGGCTCCTATGTTCT	1360
	C C T C C C C	
1361	CTTGGATACATCGTAGTGCTGAATTAAATAATTCAATTCC	1400
	A T G C C C	

FIGURE 3B

1401	TTCATCACAAATTACACAAATACCTTAACAAAATCTACT	1440
	C T C C C A G C G	
1441	AATCTTGGCTCTGGAACCTCTGTGCGTTAAGGACCAGGAT	1480
	C A G C	
1481	TTACAGGAGGAGATATTCTCGAAGAACCTCACCTGGCCA	1520
	C T A T	
1521	GATTTCAACCTTAAGAGTAAATATTACTGCACCATTATCA	1560
	AGC C C T C C C T T	
1561	CAAAGATATCGGGTAAGAATTGCTACGCTTCTACCACAA	1600
	T C G T A A	
1601	ATTTACAATTCCATACATCAATTGACGGAAGACCTATTAA	1640
	C G C C C C G C	
1641	TCAGGGGAATTTTCAGCAACTATGAGTAGTAGGGAGTAAT	1680
	T C C C C T C A C C C C	
1681	TTACAGTCCGGAAGCTTAGGACTGTAGGTTTACTACTC	1720
	G A C C A C C C	
1721	CGTTAACCTTCAAAATGGATCAAGTGTATTACGTTAAG	1760
	T C C T C C T C C T	
1761	TGCTCATGTCTCAATTCAAGGCAATGAAGTTATATAGAT	1800
	C G T G C T C	
1801	CGAATTGAATTGTTCCGGCAGAAGTAACCTTGAGGCAG	1840
	T G G T C T C T	
1841	AATAT 1845	
	G C	

FIGURE 3C

1	ATGGATAACAATCCGAACATCAATGAATGCATTCTTATA C C A C A C	40
41	ATTGTTAACGTAACCCCTGAAGTAGAAGTATTAGGTGGAGA C C G A T C T	80
81	AAGAATAGAAAATGGTTACACCCCAATCGATATTCCTTG C C T C T C C C	120
121	TCGCTAACGCAATTCTTTGAGTGAATTGTTCCGGTG CT G A G GC C C G C G A	160
161	CTGGATTGTTAGGACTAGTTGATATAATATGGGAAT G C TC C C C C T	200
201	TTTGGTCCCTCTCAATGGGACGCATTCTTGTACAAATT C A T C G G	240
241	GAACAGTTAACCAAGAATAGAAGAATTGCTAGGA G G C G G C G C	280
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTA G C G G T G C	320
321	TCAAATTACGCAGAACATCTTTAGAGAGTGGGAAGCAGAT C C T GAGC C C	360
361	CCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAAT C TC CC C G A	400
401	TCAATGACATGAACAGTGCCTTACAACCGCTATTCTCT C C T G C A C AT	440
441	TTTGCAGTCAAAATTATCAAGTTCTCTTTATCAGTA G C C G C C C G C G C G	480
481	TATGTTCAAGCTGCAAATTACATTATCAGTTTGAGAG C A T C T CC CAGC GC TC	520
521	ATGTTTCAGTGTGACAAAGGTGGGATTTGATGCCGC C AGC G C T	560
561	GACTATCAATAGTCGTTATAATGATTAACTAGGCTTATT A C C C CC T G	600
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG A C C CC C T T C T	640
641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT C G G C T T A	680

FIGURE 4A

681	AAGGTATAATCAATTAGAAGAGAATTAAACACTAACTGTA	720
	T A C C G C G G C C A T	
721	TTAGATATCGTGCCTCTGTTCCCGAATTATGATAGTAGAA	760
	G C T GT C C C CTCC	
761	GATATCCAATT CGAACAGTTCCCAATTAAACAAGAGAAAT	800
	CC C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAATTGATGGTAGTTT	840
	C T TC T G C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC	880
	T T T C A T C G CTCC C C	
881	CACATTTGATGGATATACTTAACAGTATAACCATCTATAAC	920
	C C CT G C T C	
921	GGATGCTCATAGGGTTATTATTATTGGTCAGGGCATCAA	960
	C C A AG G C T A C G G	
961	ATAATGGCTCTCCTGTAGGGTTTCGGGCCAGAATTCA	1000
	C C A T A CAGC C G T	
1001	CTTTTCCGCTATATGGAACTATGGAAATGCAGCTCCACA	1040
	C T C C C	
1041	ACAACGTATTGTTGCTCAACTAGGTCAAGGGCGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTAATATAG	1120
	C G T C G C C C C	
1121	GGATAAATAATCAACAACTATCTGTTCTTGACGGACAGA	1160
	T C C C G T C A	
1161	ATTTGCTTATGGAACCTCCTCAAATTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAGCGGAACGGTAGATTCGCTGGATGAAATAC	1240
	G C T CT C C	
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C CTC	
1281	TCATCGATTAAGCCATGTTCAATGTTGTTCAAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	
1361	CTTGGATACATCGTAGTGCTGAATTAAATAATTGCA	1400
	C G C C C C	

FIGURE 4B

1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTAATGGTTCTGTAATTCAGGACCAGGATTAC	1480
	C C C C	C
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGAAATAA	1520
	A C C C C C	
1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
1561	CCATCGACATCTACCAAGATATCGAGTTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAACCCCCGATTCACCTAACGTTAATTGGGTAA	1640
	G T	
1641	TTCATCCATTTCCTACAATACAGTACCCAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTGGTTATTTG	1720
	C G C C C C	
1721	AAAGTGCCAATGCTTTACATCTCATTAGGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTAGTGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTGAAATTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCAGGTGAATGC	1880
	A TGCG	
1881	GCTGTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT	1920
	CTGT ACCTCTACA C AGCT G ACTC G CA TG	
1921	G 1921	

FIGURE 4C

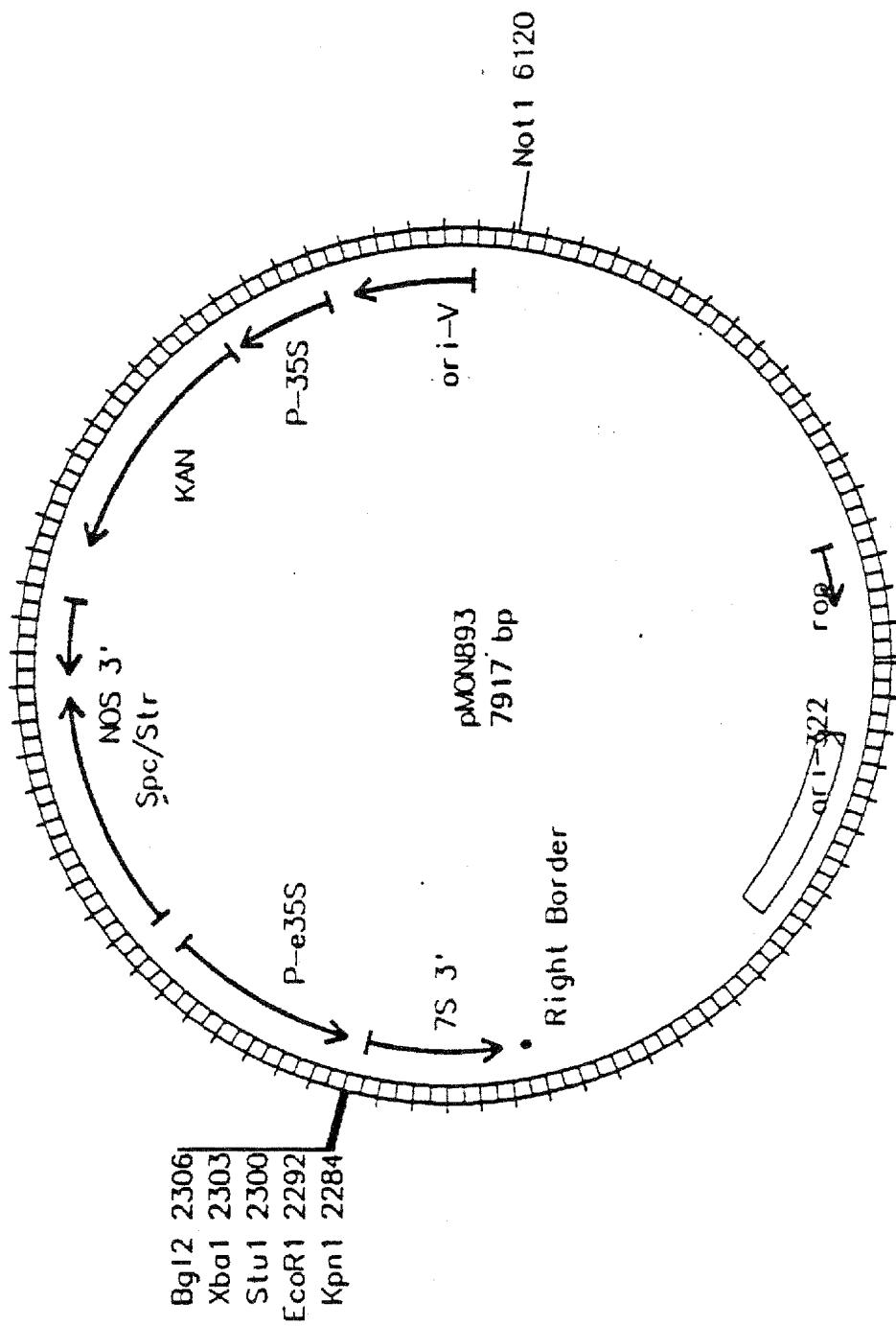


FIGURE 5

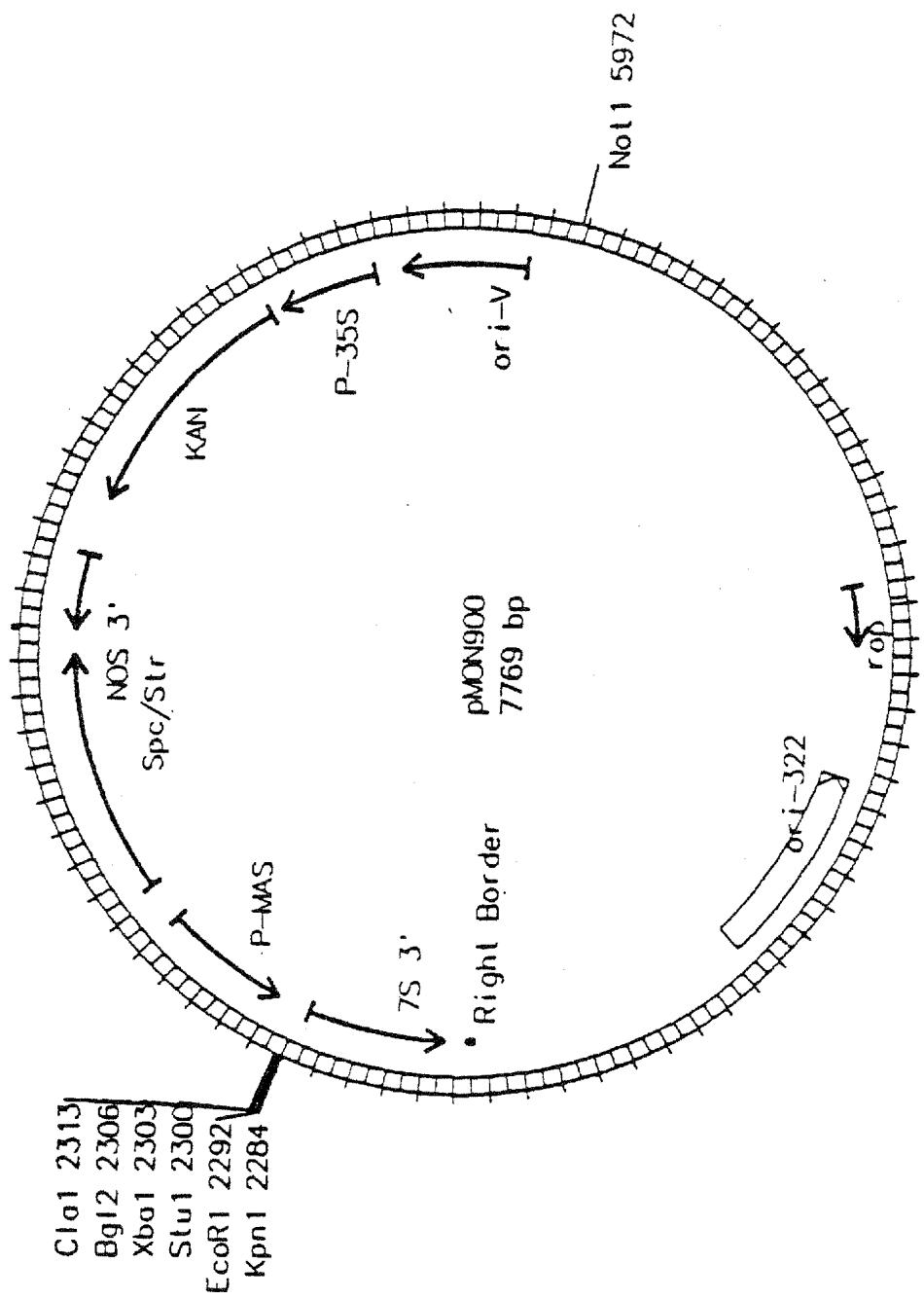


FIGURE 6

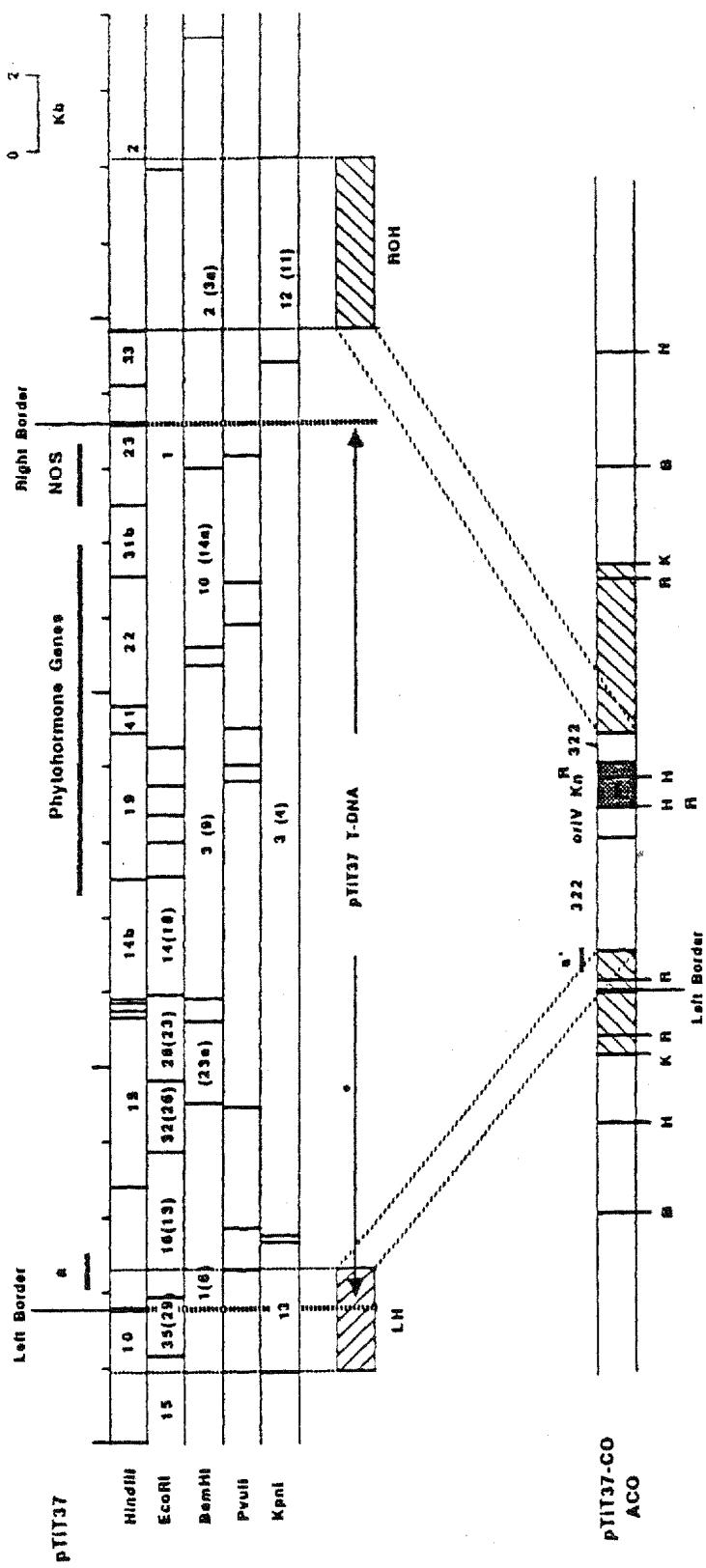


FIGURE 7

1	GAAAGAATAGAAACTGGTTACACCCCAATCGATATTCCT	40
	ATGGCC T C T C C C	
41	TGTCGCTAACGCAATTCTTTGAGTGAATTGTTCCCGG	80
	CT G A G GC C C G C G A	
81	TGCTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAA	120
	G C T C C C C C T	
121	ATTTTGTCCTCTCAATGGGACGCATTCTTGACAAA	160
	C A T C G G	
161	TTGAACAGTTAACCAAGAACATAGAACAGAACATTGCTAG	200
	G G C G G C G C	
201	GAACCAAGCCATTCTAGATTAGAACAGAACATCTT	240
	G C G G T G C	
241	TATCAAATTACGCAGAACATCTTAGAGAGATGGGAAGCAG	280
	C C T GAGC C C	
281	ATCCTACTAATCCAGCATTAAGAGAACAGAGATGCGTATTCA	320
	C TC CC C G A	
321	ATTCAATGACATGAACAGTGCCCTTACAACCGCTATTCT	360
	C C T G C A C A	
361	CTTTTGCAAGTTCAAAATTATCAAGTTCTCTTTATCAG	400
	T G C C G C C C G C	
401	TATATGTTCAAGCTGCAAATTACATTATCAGTTTGAG	440
	G C A T C T CC CAGC GC TC	
441	AGATGTTCAAGTGTGGACAAAGGTGGGATTGATGCC	480
	C AGC G C T	
481	GCGACTATCAATAGTCGTTATAATGATTTAACAGGCTTA	520
	A C C C C C C T G	
521	TTGGCAACTATACAGATTATGCTGTACGCTGGTACAATAC	560
	A C C C C C T T C	
561	GGGATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGG	600
	T C G G C T T	
601	GTAAGGTATAATCAATTAGAACAGAACACTAACTG	640
	A T A C C G C G G C C A	
641	TATTAGATATCGTTGCTCTGTTCCCAGATTATGATAGTAG	680
	T G C T GT C C T C	

FIGURE 8A

681	AAGATATCCAATTGAAACAGTTCCCAATTAAACAAGAGAA	720
	CC C T C T G C T C	
721	ATTTATACAAACCCAGTATTAGAAAATTTGATGGTAGTT	760
	C T TC T G C C C	
761	TTCGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAG	800
	C T T C A T C G CTCC C	
801	TCCACATTTGATGGATATACTTAACAGTATAACCATCTAT	840
	C C C CT G C T C	
841	ACGGATGCTCATAGGGTTATTATTATTGGTCAGGGCATC	880
	C C A AG G C T A C	
881	AAATAATGGCTTCTCCTGTAGGGTTTCGGGCCAGAATT	920
	G C C A T A CAGC C G	
921	CACTTTCCGCTATATGGAACTATGGAAATGCAGCTCCA	960
	T C T C C C	
961	CAACAACGTATTGTTGCTCAACTAGGTCAAGGGCGTGTATA	1000
	C T C C	
1001	GAACATTATCGTCCACTTTATATAGAACCTTTAATAT	1040
	C G T C G C C C	
1041	AGGGATAAATAATCAACAACTATCTGTTCTTGACGGGACA	1080
	C T C C C G T C A	
1081	GAATTTGCTTATGGAACCTCCTCAAATTGCCATCCGCTG	1120
	G C C T T C	
1121	TATACAGAAAAAGCGGAACGGTAGATTGCTGGATGAAAT	1160
	T G C T C T C	
1161	ACCGCCACAGAATAACAACGTGCCACCTAGGCAAGGATT	1200
	C A C T C C	
1201	AGTCATCGATTAAGCCATGTTCAATGTTGTTCAAGGCT	1240
	TCC CA G G C G C C A	
1241	TTAGTAATAGTAGTGTAAAGTATAATAAGAGCTCCTATGTT	1280
	C C C TCC G C C C	
1281	CTCTTGGATAACATCGTAGTGCTGAATTAAATAATTAAATT	1320
	C G C C C C C	
1321	GCATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAA	1360
	C	
1361	ACTTTCTTTTAATGGTTCTGTAATTTCAGGACCAGGATT	1400
	C C C C	

FIGURE 8B

1401	TACTGGTGGGGACTTAGTTAGATTAAATAGTAGTGAAAT	1440
	C A C C C C C C	
1441	AACATTCAAGAATAGAGGGTATATTGAAGTTCCAATTCACT	1480
1481	TCCCACATCGACATCTACCAGATATCGAGTTCGTGTACGGTA	1520
	C A GA	
1521	TGCTTCTGTAACCCCGATTCACCTAACGTTAATTGGGGT	1560
	G T	
1561	AATTCACTCCATTTTCCAATACAGTACCACTACAGCTA	1600
	C C T	
1601	CGTCATTAGATAATCTACAATCAAGTGATTTGGTTATT	1640
	C C G C C C C C	
1641	TGAAAGTGCCAATGCTTTACATCTCATTAGGTAATATA	1680
	C C C C	
1681	G TAGGTGTTAGAAATTTAGTGGGACTGCAGGAGTGATAA	1720
	G C T	
1721	TAGACAGATTGAATTTATTCCAGTTACTGCAACACTCGA	1760
	C C G C	
1761	GGCTGAA 1767	
	G	

FIGURE 8C

1	ATGGATAACAATCCGAACATCAATGAATGCATTCTTATA	40
	C C A C A C	
41	ATTGTTAACGTAACCCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAAAATGGTTACACCCCCAATCGATATTCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTCTTTGAGTGAATTGTTCCCGGTG	160
	CT G A G G C C C G C G A	
161	CTGGATTGTGTTAGGACTAGTTGATATAATATGGGAAT	200
	G C T C C C C C T	
201	TTTGTCCTCTCAATGGGACGCATTCTTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAACCAACAAAGAATAGAAGAATTGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTA	320
	G C G G T G C	
321	TCAAATTACGCAGAACATCTTAGAGAGTGGGAAGCAGAT	360
	C C T G A C G C	
361	CCTACTAACCGCATTAAGAGAAGAGATGCGTATTCAAT	400
	C T C C C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCTCT	440
	C C T G C A C AT	
441	TTTGCAGTTCAAAATTATCAAGTTCCCTTTTATCAGTA	480
	G C C G C C C G C G C G	
481	TATGTTCAAGCTGCAAATTACATTTATCAGTTTGAGAG	520
	C A T C T C C C A G C G C T C	
521	ATGTTTCAGTGTGGACAAAGGTGGGATTTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C C T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C C C T T C T	
641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT	680
	C G G C T T A	

FIGURE 9A

681	AAGGTATAATCAATTAGAAGAGAATTAACACTAACTGTA T A C C G C G G C C A T	720
721	TTAGATATCGTTGCTCTGTTCCCGAATTATGATAGTAGAA G C T GT C C CTCC	760
761	GATATCCAATTGAAACAGTTCCCAATTAAACAAGAGAAAT CC C T C T G C T C	800
801	TTATACAAACCCAGTATTAGAAAATTGATGGTAGTTT C T TC T G C C C C C	840
841	CGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC T T C A T C G C T C C	880
881	CACATTGATGGATATACTTAACAGTATAACCATCTATAC C C T G C T C	920
921	GGATGCTCATAGGGTTATTATTATTGGTCAGGGCATCAA C C A AG G C T A C G	960
961	ATAATGGCTTCTCCTGTAGGGTTTCGGGCCAGAATTCA C C A T A CAGC C G T	1000
1001	CTTTTCCGCTATATGGAACATATGGAAATGCAGCTCCACA C T C C C	1040
1041	ACAACGTATTGTTGCTCAACTAGGTCAAGGGCGTGTATAGA C T C C	1080
1081	ACATTATCGTCCACTTATATAGAACGCTTTAATATAG C G T C G C C C C	1120
1121	GGATAAATAATCAACAACTATCTGTTCTGACGGGACAGA T C C G T C A	1160
1161	ATTTGCTTATGGAACCTCCTCAAATTGCCCCTCGCTGTA G C C T T C T	1200
1201	TACAGAAAAAGCGGAACGGTAGATTGCTGGATGAAATAC G C T C T C C	1240
1241	CGCCACAGAATAACAAACGTGCCACCTAGGCAAGGATTTAG A C T C CTC	1280
1281	TCATCGATTAAGCCATGTTCAATGTTGTTCAAGGCTTT C CA G G C G C C C A C	1320
1321	AGTAATAGTAGTGTAAAGTATAATAAGAGCTCCTATGTTCT C C TCC G C C C	1360
1361	CTTGGATACATCGTAGTGCTGAATTAAATAATTGCG C G C C C C C	1400

FIGURE 9B

1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTTAATGGTTCTGTAATTCAAGGACCAGGATTTA	1480
	C C C C	C
1481	CTGGTGGGGACTTAGTTAGATAATAGTAGTGAAATAA	1520
	A C C C C C	
1521	CATTCAGAACATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
1561	CCATCGACATCTACCAAGATATCGAGTCGTGTACGGTATG	1600
	C A	GA
1601	CTTCTGTAACCCCGATTCACCTAACGTTAATTGGGTAA	1640
	G T	
1641	TTCATCCATTTCACAATACAGTACCAAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTGGTTATTTG	1720
	C G C C C C	
1721	AAAGTGCCAATGCTTTACATCTCATTAGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C	T C
1801	GACAGATTGAAATTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
1881	GCTGTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT	1920
1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTAGTTA	1960
1961	CGTATTATCGGATGAATTTGTCTGGATGAAAAGCGAGA	2000
2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
2041	GAACGCAATTACTCCAAGATTCAAATTCAAAGACATTA	2080
2081	ATAGGCAACCAGAACGTGGGTGGGCGGAAGTACAGGGAT	2120

FIGURE 9C

2121	TACCATCCAAGGAGGGATGACGTATTAAAGAAAATTAC	2160
2161	GTCACACTATCAGGTACCTTGATGAGTGCTATCCAACAT	2200
2201	ATTTGTATCAAAAAATCGATGAATCAAAATTAAAAGCCTT	2240
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA	2280
2281	GACTTAGAAATCTATTTAATTGCTACAATGCAAAACATG	2320
2321	AAACAGTAAATGTGCCAGGTACGGGTTCCCTATGCCGCT	2360
2361	TTCAGCCAAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
2441	GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATTCGCA	2480
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
2521	AATGAGGACCTAGGTGTATGGGTGATCTTAAGATTAAGA	2560
2561	CGCAAGATGGCACGCAAGACTAGGGATCTAGAGTTCT	2600
2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
2641	AAAAGAGCGGAGAAAAATGGAGAGACAAACGTGAAAAT	2680
2681	TGGAATGGAAACAAATATCGTTATAAGAGGGCAAAAGA	2720
2721	ATCTGTAGATGCTTATTTGTAAACTCTCAATATGATCAA	2760
2761	TTACAAGCGGATACGAATATTGCCATGATTGCGGCAG	2800
2801	ATAAACGTGTTCATAGCATTGAGAAGCTTATCTGCCTGA	2840

FIGURE 9D

2841	GCTGTCTGTGATTCCGGGTGTCATGCAGCTATTTTGAA	2880
2881	GAATTAGAAGGGCGTATTTCACTGCATTCTCCCTATATG	2920
2921	ATGCGAGAAATGTCATTAAAAATGGTGATTTAATAATGG	2960
2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000
3001	GAACAAAACAACCAACGTTGGTCCTTGTGTTCCGGAAAT	3040
3041	GGGAAGCAGAAGTGTACAAAGAAGTTCGTGTCTGTCCGGG	3080
3081	TGTTGGCTATATCCTTCGTGTCACAGCGTACAAGGAGGG	3120
3121	TATGGAGAAGGTTGCGTAACCATTGAGATCGAGAACAA	3160
3161	ATACAGACGAACTGAAGTTAGCAACTGCGTAGAACAGAGGA	3200
3201	AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT	3240
3241	GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCGTA	3280
3281	ATCGAGGATATAACGAAGCTCCTCCGTACCGAGCTGATT	3320
3321	TGCGTCAGTCTATGAAGAAAATCGTATACAGATGGACGA	3360
3361	AGAGAGAATCCTTGTGAATTAAACAGAGGGTATAGGGATT	3400
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
3481	GAAACGGAAGGAACATTTATCGTGGACAGCGTGGATTAC	3520
3521	TCCTTATGGAGGAA 3534	

FIGURE 9E

1	ATGGATAACAATCCGAACATCAATGAATGCATTCTTATA	40
	C C A C A C	
41	ATTGTTTAAGTAACCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAAAATGGTTACACCCCCAATCGATATTCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTCTTTGAGTGAATTGTTCCCGGTG	160
	CT G A G G C C C G C G A	
161	CTGGATTGTGTTAGGACTAGTTGATATAATGGGAAAT	200
	G C T C C C C C T	
201	TTTGTCCTCTCAATGGGACGCATTCTTGTACAAATT	240
	C A T C G G G	
241	GAACAGTTAATTAACCAAAGAATAGAAGAATTGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTA	320
	G C G G T G C	
321	TCAAATTACGCAGAATCTTTAGAGAGTGGGAAGCAGAT	360
	C C T G A G C C	
361	CCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAAT	400
	C T C C C G A	
401	TCAATGACATGAACAGTGCCCTACAACCGCTATTCTCT	440
	C C T G C A C AT	
441	TTTGCAGTTCAAAATTATCAAGTTCTCTTTATCAGTA	480
	G C C G C C C G G G	
481	TATGTTCAAGCTGCAAATTACATTATCAGTTGAGAG	520
	C A T C T C C C A G C G C T C	
521	ATGTTTCAGTGTGGACAAAGGTGGGATTTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTAACTAGGCTTATT	600
	A C C C C C C T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C C C T T C T	

FIGURE 10A

641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT C G G C T T A	680
681	AAGGTATAATCAATTAGAAGAGAATTAACACTAACTGTA T A C C G C G G C C A T	720
721	TTAGATATCGTTGCTCTGTTCCCGAATTATGATAGTAGAA G C T GT C C CTCC	760
761	GATATCCAATTGAAACAGTTCCCAATTAAACAAGAGAAAT CC C T C T G C T C	800
801	TTATACAAACCCAGTATTAGAAAATTTGATGGTAGTTT C T TC T G C C C C C	840
841	CGAGGCTCGGCTCAGGGCATAGAAAGTATTAGGAGTC T T T C A T C G C T C C	880
881	CACATTGATGGATATACTTAACAGTATAACCATCTATA C C T G C T C	920
921	GGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATCAA C C A AG G C T A C G	960
961	ATAATGGCTTCTCCTGTAGGGTTTCGGGCCAGAATTCA C C A T A C A G C C G T	1000
1001	CTTTTCCGCTATATGGAACTATGGAAATGCAGCTCCACA C T C C	1040
1041	ACAACGTATTGTTGCTCAACTAGGTCAAGGGCGTGTATAGA C T C C	1080
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTAATATAG C G T C G C C C C	1120
1121	GGATAAAATAATCAACAACTATCTGTTCTTGACGGGACAGA T C C G T C A	1160
1161	ATTTGCTTATGGAACCTCCTCAAATTGCCATCCGCTGTA G C C T T C T	1200
1201	TACAGAAAAAGCGGAACGGTAGATTGCTGGATGAAATAC G C T C T C C	1240
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG A C T C CTC	1280
1281	TCATCGATTAAGCCATGTTCAATGTTGTTCAAGGCTTT C CA G G C G C C C A C	1320
1321	AGTAATAGTAGTGTAAAGTATAATAAGAGCTCCTATGTTCT C C T C G C C C	1360

FIGURE 10B

1361	CTTGGATACATCGTAGTGCIGAATTAAATAATATAATTGC C G C C C C C	1400
1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC C	1440
1441	TTTCTTTTAATGGTTCTGTAAATTCAGGACCAGGATTAA C C C C C	1480
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAAATAA A C C C C C C	1520
1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
1561	CCATCGACATCTACCAAGATATCGAGTCGTGTACGGTATG C A GA	1600
1601	CTTCTGTAACCCCGATTCACCTAACGTTAATTGGGTAA G T	1640
1641	TTCATCCATTTCACAATACAGTACCAAGCTACAGCTACG C C T C	1680
1681	TCATTAGATAATCTACAATCAAGTGATTTGGTTATTTG C G C C C C C	1720
1721	AAAGTGCCAATGCTTTACATCTTCATTAGGTAAATATAGT C C C C C	1760
1761	AGGTGTTAGAAATTTAGTGGACTGCAGGAGTGATAATA G C T C	1800
1801	GACAGATTGAAATTATTCCAGTTACTGCAACACTCGAGG C G C	1840
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
1881	GCTGTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT G C C C G C	1920
1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTAGTTA G C G G	1960
1961	CGTATTATCGGATGAATTTGTCTGGATGAAAAGCGAGA C CC CAGC G C	2000
2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
2041	GAACGCAATTACTCCAAGATTCAAATTCAAAGACATTA	2080

FIGURE 10C

2081	ATAGGCAACCAGAACGTGGGTGGGCAGAAGTACAGGGAT	2120
2121	TACCATCCAAGGAGGGATGACGTATTTAAAGAAAATTAC G T C G C G G C	2160
2161	GTCACACTATCAGGTACCTTGATGAGTGCTATCCAACAT	2200
2201	ATTTGTATCAAAAAATCGATGAATCAAATTAAAGCCTT CC C C G G C G C G G	2240
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAAGTCAA	2280
2281	GACTTAGAAATCTATTTAATTGCTACAATGCAAAACATG C C G CC C C	2320
2321	AAACAGTAAATGTGCCAGGTACGGGTTCCCTATGGCCGCT	2360
2361	TTCAGCCCAGTCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
2441	GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATTCGCA	2480
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
2521	AATGAGGACCTAGGTATGGGTGATTTAAGATTAAGA	2560
2561	CGCAAGATGGCACGCAAGACTAGGAAATCTAGAGTTCT	2600
2601	CGAAGAGAAACCATTAGTAGGAGAACGCTAGCTCGTGTG	2640
2641	AAAAGAGCGGAGAAAAATGGAGAGACAAACGTGAAAAT G G	2680
2681	TGGAATGGAAACAAATATCGTTATAAGAGGGCAAAAGA G C C C C	2720
2721	ATCTGTAGATGCTTATTTGAACTCTCAATATGATCAA	2760
2761	TTACAAGCGGATACGAATATTGCCATGATTGCGGCAG	2800

FIGURE 10D

2801	ATAAACGTGTTCATAGCATTGAGAAGCTTATCTGCCTGA	2840
2841	GCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTGAA	2880
2881	GAATTAGAAGGGCGTATTTCACTGCATTCTCCCTATATG C C	2920
2921	ATGCGAGAAAATGTCATTAATAATGGTGTATTTAATAATGG C C C G C C C C	2960
2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000
3001	GAACAAAACAACCAACGTTCGTCCTGTTGTTCCGGAAAT	3040
3041	GGGAAGCAGAAGTGTACAAGAAGTTCGTGTCTGTCGGG	3080
3081	TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGGAA	3120
3121	TATGGAGAAGGTTGCGTAACCATTGAGATCGAGAACAA	3160
3161	ATACAGACGAACTGAAGTTAGCAACTGCGTAGAAGAGGA	3200
3201	AATCTATCAAATAACACGTAACGTGTAATGATTATACT	3240
3241	GTAAATCAAGAAGAATAACGGAGGTGCGTACACTTCGTA	3280
3281	ATCGAGGATATAACGAAGCTCCTCCGTACCGACTGATTA	3320
3321	TGCGTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA	3360
3361	AGAGAGAATCCTTGTGAATTAAACAGAGGGTATAGGGATT	3400
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
3481	GAAACGGAAGGAACATTTATCGTGGACAGCGTGGATTAC	3520
3521	TCCTTATGGAGGAA 3534	

FIGURE 10E

1	ATGGATAACAATCCGAACATCAATGAATGCATTCTTATA C C A C A C	40
41	ATTGTTAACGTAACCCCTGAAGTAGAAGTATTAGGTGGAGA C C G A T C T	80
81	AAGAATAGAAACTGGTTACACCCCAATCGATATTCCTTG C C T C T C C C	120
121	TCGCTAACGCAATTCTTTGAGTGAATTTGTTCCCGGTG CT G A G GC C C G C G A	160
161	CTGGATTGTTAGGACTAGTTGATATAATATGGGAAT G C TC C C C C T	200
201	TTTGTCCTCTCAATGGACGCATTCTTGACAAATT C A T C G G	240
241	GAACAGTTAACCAAGAATAGAAGAATTGCTAGGA G G C G G C G C	280
281	ACCAAGCCATTCTAGATTAGAAGGACTAAGCAATCTTA G C G G T G C	320
321	TCAAATTACGCAGAACATCTTTAGAGAGTGGGAAGCAGAT C C T GAGC C C	360
361	CCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAAT C TC CC C G A	400
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCTCT C T G C A C AT	440
441	TTTGCAGTTCAAAATTATCAAGTTCTCTTATCAGTA G C C G C C C G C G C G	480
481	TATGTTCAAGCTGCAAATTACATTATCAGTTTGAGAG C A T C T CC CAGC GC TC	520
521	ATGTTTCAGTGGACAAAGGTGGGATTTGATGCCGC C AGC G C T	560
561	GACTATCAATAGTCGTTATAATGATTTAACAGGCTTATT A C C C CC T G	600
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG A C C CC C T T C T	640
641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT C G G C T T A	680

FIGURE 11A

681	AAGGTATAATCAATTAGAAGAGAATTAACACTAAGTGA T A C C G C G G C C A T	720
721	TTAGATATCGTTGCTCTGTTCCCGAATTATGATAGTAGAA G C T GT C C CTCC	760
761	GATATCCAATTGAAACAGTTCCCAATTAAACAAGAGAAAT CC C T C T G C T C	800
801	TTATACAAACCCAGTATTAGAAAATTTGATGGTAGTTT C T TC T G C C C C C	840
841	CGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC T T T C A T C G CTCC C C	880
881	CACATTTGATGGATATACTAACAGTATAACCATCTATAC C C CT G C T C	920
921	GGATGCTCATAGGGTTATTATTATTGGTCAGGGCATCAA C C A AG G C T A C G	960
961	ATAATGGCTTCTCCTGTAGGGTTTCGGGCCAGAATTCA C C A T A CAGC C G T	1000
1001	CTTTTCCGCTATATGAACTATGGAAATGCAGCTCCACA C T C C C	1040
1041	ACAACGTATTGTTGCTCAACTAGGTCAAGGGCGTGTATAGA C T C C	1080
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTAATATAG C G T C G C C C C	1120
1121	GGATAAATAATCAACAACTATCTGTTCTGACGGGACAGA T C C C G T C A	1160
1161	ATTTGCTTATGAAACCTCCTCAAATTGCCATCCGCTGTA G C C T T C T	1200
1201	TACAGAAAAAGCGGAACCGGTAGATTGCTGGATGAAATAC G C T CT C C	1240
1241	CGCCACAGAATAACAAACGTGCCACCTAGGCAAGGATTTAG A C T C CTC	1280
1281	TCATCGATTAAGCCATGTTCAATGTTGCTCAGGCTTT C CA G G C G C C C A C	1320
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT C C TCC G C C C	1360
1361	CTTGGATACATCGTAGTGCTGAATTAAATAATTGAC C G C C C C C	1400

FIGURE 11B

1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTAATGGTCTGTAATTCAAGGACCAGGATTAC	1480
	C C C C	C
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGAAATAA	1520
	A C C C C C	
1521	CATTCAGAACAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
1561	CCATCGACATCTACCAAGATATCGAGTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAACCCCGATTCACCTCAACGTTAATTGGGTAA	1640
	G T	
1641	TTCATCCATTTTCCAATACAGTACCAAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTGGTTATTTG	1720
	C G C C C C	
1721	AAAGTGCCAATGCTTTACATCTTCATTAGGTAAATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTTAGTGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTGAATTTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCCGTGAATGC	1880
	G C C T G C T C	
1881	GCTGTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT	1920
	C C C C T G T CT G T C	
1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTAGTTA	1960
	T T C C C C C G C	
1961	CGTATTTATCGGATGAATTTGTCTGGATGAAAAGCGAGA	2000
	C CC TAGC G C C C C G T	
2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
	C C T C C T C C	
2041	GAACGCAATTACTCCAAGATTCAAATTCAAAGACATTA	2080
	GA G C CT G C C C C	
2081	ATAGGCAACCAGAACGTGGTGGGGCGGAAGTACAGGGAT	2120
	C G T T C C	

FIGURE 11C

2121	TACCATCCAAGGAGGGATGACGTATTTAAGAAAATTAC C C C T G C G G C	2160
2161	GTCACACTATCAGGTACCTTGATGAGTGCTATCCAACAT C C C A T C C C T C	2200
2201	ATTTGTATCAAAAAATCGATGAATCAAATTAAAAGCCTT C C G G G C C C	2240
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTC C A G C T C C C C	2280
2281	GACTTAGAAATCTATTTAACGCTACAATGC C T C CG CA G C G C	2320
2321	AAACAGTAAATGTGCCAGGTACGGGTTCTTATGGCGCT G C G C T C C A	2360
2361	TTCAGCCCAGTCCAATCGGAAAGTGTGGAGAGCCGAAT T TC C T G T C	2400
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT A T G G C	2440
2441	GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATCATTGCA C C C C G C T	2480
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA C G C G T C G	2520
2521	AATGAGGACCTAGGTGTATGGGTGATCTTAAGATTAAGA C A C C C C	2560
2561	CGCAAGATGGCACGCAAGACTAGGGATCTAGAGTTCT C C A T C C T	2600
2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG G C T T C	2640
2641	AAAAGAGCGGAGAAAAATGGAGAGACAAACGTAAAAAT G A G G G G C	2680
2681	TGGAATGGGAAACAAATATCGTTATAAAGAGGCAAAAGA C T C C G C	2720
2721	ATCTGTAGATGCTTATTTGTAACCTCTCAATATGATCAA G C G G C G C G	2760
2761	TTACAAGCGGATACGAATATTGCCATGATTGCGGCAG G C C C C C C C	2800
2801	ATAAACGTGTTCATAGCATTGAGAAGCTTATCTGCCTGA C G C T G C T	2840

FIGURE 11D

2841	GCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTGAA	2880
	T C C T G C T C C C G	
2881	GAATTAGAACGGCGTATTTCACTGCATTCTCCCTATATG	2920
	C T G A C T C T G C	
2921	ATGCGAGAAATGTCATTAAAATGGTGATTTAATAATGG	2960
	C C C G C C C C	
2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000
	C CAG T T G C G G G	
3001	GAACAAAACAACCAACGTTCGGTCTTGTGTTCCGGAAT	3040
	G T G C G G T G	
3041	GGGAAGCAGAACAGTGTACAAAGAAGTCGTGTCTGCCGG	3080
	T C G A A A A	
3081	TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGG	3120
	A A C T C G C T	
3121	TATGGAGAACGGTTGCGTAACCATTATGAGATCGAGAAC	3160
	C T G G C C	
3161	ATACAGACGAACTGAAGTTAGCAACTGCGTAGAACAGAGGA	3200
	C C G T C T C G A	
3201	AATCTATCCAAATAACACCGTAACGTGTAATGATTATACT	3240
	C C C T T C C C C	
3241	GTAAATCAAGAAGAACGGGAGGTGCGTACACTTCTCGTA	3280
	G G G C C AGC	
3281	ATCGAGGATATAACGAAGCTCCTTCCGTACCAAGCTGATT	3320
	CA T C T T C	
3321	TGCGTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA	3360
	C C G C G G C C CA	
3361	AGAGAGAACCTTGTGAATTAAACAGAGGGTATAGGGATT	3400
	C T C C G C T C C	
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
	A T C T C G G C T	
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
	G T T G C A G C C T	
3481	GAAACGGAAGGAACATTTATCGTGGACAGCGTGGATTAC	3520
	C G C C GC T	
3521	TCCTTATGGAGGAA	3534
	T G	

FIGURE 11E

1	ATGACTGCAGATAATAATACGGAAAGCACTAGATAGCTCTA C C C C C C C T	40
41	CAACAAAAGATGTCATTCAAAAAGGCATTCGGTAGTAGG C T G T C G G T C T G	80
81	TGATCTCCTAGGCCTAGTAGGTTCCGTTGGTGGAGCG A C T G G T A T C C C	120
121	CTTGTTCGTTTATACAAACTTTAAACTATTTGGC C GAGC C C C C	160
161	CAAGTGAAGACCCGTGGAAGGCTTTATGGAACAAGTAGA C G T A A C G T	200
201	AGCATTGATGGATCAGAAAATAGCTGATTATGCAAAAAAT TC T G T A C G C	240
241	AAAGCTCTTGCAGAGTTACAGGGCCTCAAAATAATGTCG G T G AC C G C G	280
281	AAGATTATGTGAGTGCATTGAGTTATGGCAAAAAATCC G C C TCCAGC G G C	320
321	TGTGAGTTCACGAAATCCACATAGCCAGGGCGGATAAGA T C CA T C A TA C	360
361	GAGCTTTCTCAAGCAGAAAGTCATTCGTAATTCAA T C C TCC C CA A C	400
401	TGCCTTCGTTGCAATTCTGGATACGAGGTTCTATTCT AGC T C C T T C	440
441	AACAAACATATGCACAAGCTGCCAACACACATTATTTA C T C T C C G C C	480
481	CTAAAAGACGCTCAAATTATGGAGAAGAATGGGGATACG T G C G	520
521	AAAAAGAAGATATTGCTGAATTATAAAAGACAACATAA G G C G C C G C T T	560
561	ACTTACGCAAGAATATACTGACCATTGTGTCAAATGGTAT G C C G C C G	600
601	AATGTTGGATTAGATAAAATTAAAGAGGTTCATCTTATGAAT C TC C GC C C T C C G	640
641	CTTGGGTAAACTTAACCGTTATCGCAGAGAGATGACATT G C A A C A G C	680

FIGURE 12A

681	AACAGTATTAGATTTAATTGCACTATTCATTTGAT	720
	G T GC C C T C C C C	
721	GTTCGGCTATAACCCAAAAGAAGTTAAAACCGAATTAACAA	760
	GA A C G G T GC T C	
761	GAGACGTTTAACAGATCCAATTGTCGGAGTCAACAACCT	800
	GC C T C T	
801	TAGGGGCTATGGAACAAACCTCTCTAATATAGAAAATTAT	840
	T T AGC C C C	
841	ATTCGAAAACCACATCTATTGACTATCTGCATAGAATTCA	880
	A G C C T C	
881	AATTCACACGCGTTCCAACCAGGATATTATGAAATGAA	920
	C AA T C T C	
921	CTCTTCATTATTGGTCCGGTAATTATGTTCAACTAGA	960
	C C C C C C	
961	CCAAGCATAGGATCAAATGATATAATCACATCTCCATTCT	1000
	T T C C C C	
1001	ATGGAAATAAATCCAGTGAACCTGTACAAAATTAGAATT	1040
	T C G G G CCT T G	
1041	TAATGGAGAAAAAGTCTATAGAGCCGTAGCAAATACAAAT	1080
	C C C G C C C	
1081	CTTGCGGTCTGGCCGTCCGCTGTATATTCAAGGTGTTACAA	1120
	C T G A A T C C C	
1121	AAGTGGAAATTAGCCAATATAATGATCAAACAGATGAAGC	1160
	G G T G C G C G	
1161	AAGTACACAAACGTACGACTCAAAAAGAAATGTTGGCGCG	1200
	C C C G T C T C A	
1201	GTCAGCTGGGATTCTATCGATCAATTGCCTCCAGAAACAA	1240
	TCT C C	
1241	CAGATGAACCTCTAGAAAAGGGATATGCCATCAACTCAA	1280
	C AT G G C C C T	
1281	TTATGTAATGTGCTTTAATGCAGGGTAGTAGAGGAACA	1320
	C G C G A T C C G C	
1321	ATCCCAGTGTAACTGGACACATAAAAGTGTAGACTTT	1360
	T G C C GTCC G C	
1361	TTAACATGATTGATTGAAAAAAATTACACAACTTCCGTT	1400
	C C AGC G G C T C	

FIGURE 12B

1401	AGTAAAGGCATATAAGTTACAATCTGGTGCTTCGTTGTC	1440
	G G A C C C G	
1441	GCAGGTCCCTAGGTTACAGGAGGAGATCATTCAATGCA	1480
	C A C T T C C G	
1481	CAGAAAATGGAAGTGC GGCAACTATTTACGTTACACCGGA	1520
	G C C C A T C G T	
1521	TGTGTCGTACTCTCAAAAATATCGAGCTAGAATTCAATTAT	1560
	T G G C A G A C T C	
1561	GCTTCTACATCTCAGATAACATTACACTCAGTTAGACG	1600
	A CAGC C C C C G T	
1601	GGGCACCATTAAATCAATACTATTCGATAAAACGATAAA	1640
	A C C C G T C T C G C C	
1641	TAAAGGAGACACATTAACGTATAATTCAATTAAATTAGCA	1680
	C T T C C A C AGC C C G	
1681	AGTTTCAGCACACCATTCAATTATCAGGGATAACTTAC	1720
	T C C C C T C T	
1721	AAATAGGCGTCACAGGATTAAGTGCTGGAGATAAAGTTA	1760
	G C C T C C C C C	
1761	TATAGACAAAATTGAATTATTCCAGTGAAT	1791
	C C G G C C C	

FIGURE 12C

1	ATG	AATAATGTATTGAATAGTGGAAAGAACAACTATT	40
	GAC	C C C CTC T C C	
41	GTGATGCGTATAATGTAGTAGCCCATGATCCATTAGTT	80	
	C C A C C C G T C	C C	
81	TGAACATAAATCATTAGATACCATCCAAAAAGAATGGATG	120	
	C C GAGCC C C T T G G G		
121	GAGTGGAAAAGAACAGATCATAGTTATATGTAGCTCCTG	160	
	A C T T C CTC C C C C A		
161	TAGTCGAACTGTGTCTAGTTTTGCTAAAGAAAGTGGG	200	
	G T A C C CC T C G C		
201	GAGTCTTATTGGAAAAGGATATTGAGTGAATTATGGGGG	240	
	CTC C C C T C TCC C C T		
241	ATAATATTCCTAGGGTAGTACAATCTAATGCAAGATA	280	
	C C ATC GTCC T C C		
281	TTTAAGGGAGACAGAACAAATTCTAAATCAAAGACTTAA	320	
	C G C G T C C GC T C		
321	TACAGATAACCCTTGCTCGTGTAAATGCAGAATTGATAGGG	360	
	C T T G A A C C T G C T		
361	CTCCAAGCGAATATAAGGGATTTAATCAACAAAGTAGATA	400	
	A C TC T C C G G C		
401	ATTTTTAAACCTACTCAAAACCTGTTCTTATCAAT	440	
	C C G T A G T G C T C		
441	AACTTCTCGGTTAATACAATGCAGCAATTATTCTAAAT	480	
	C C G C T C C C C C		
481	AGATTACCCAGTCCAGATACAAGGATACCAGTTGTTAT	520	
	G T T T C C		
521	TATTACCTTATTGCACAGGCAGCCAATATGCATCTTC	560	
	TC T AC C T T C CT G		
561	TTTTATTAGAGATGTTATTCTTAATGCAGATGAATGGGGT	600	
	C C AC T C G C C C T C A		
601	ATTCAGCAGCAACATTACGTACGTATCGAGATTACCTGA	640	
	C T C T C T A G A C A C T		
641	GAAATTATAACAAGAGATTATTCTAATTATTGTATAAATAC	680	
	G C C TC T C C C C C C		

FIGURE 13A

681	GTATCAAACCGCTTAGAGGGTAAACACCCGTTACAC T G C C T A C C T A G C T	720
721	GATATGTTAGAATTAGAACATATACTGTTAAATGTAT C C T G C G C C C C T C G	760
761	TTGAATATGTATCCATTTGGTCATTGTTAAATATCAGAG G C C A G A G T C C C G C	800
801	TCTTATGGTATCTCTGGCGCTAATTTATATGCTAGCGGT C T G G C A C C C C C T C T C	840
841	AGTGGACCACAGCAGACACAATCATTTACAGCACAAA A T G A G C C T G	880
881	GGCCATTTTATATTCTCTTTCCAAGTTAATTGAAATT C G A G C T G C C C C	920
921	TATATTATCTGGTATTAGTGGTACTAGGCTTCTATTACC C T C C A G C T C G C A C C A	960
961	TTCCCTAATATTGGTGGTTACCGGGTAGTACTACA T C C A C T A C T C C C	1000
1001	ATTCAATTGAATAGTGCCAGGGTTAATTATAGCGGAGGAGT A G C C T T	1040
1041	TTCATCTGGTCTCATAGGGCGACTAATCTCAATCACAA C A G C T T A C T G C	1080
1081	TTTAATTGCAGCACGGCCTCCCTCCTTATCAACACC C T C T G A C G A G C G	1120
1121	TTGTTAGAAGTTGGCTGGATTCAAGGTACAGATCGAGAGGG G GTCC T C A G C T C A	1160
1161	CGTTGCTACCTCTACGAATTGGCAGACAGAACATCCTTCAA A C A C G C	1200
1201	ACAACTTAAAGTTAAGGTGTGGTGCTTTCAAGCCGTG C C T C C T C A C T A	1240
1241	GAAATTCAAACATTTCCCAGATTATTTATCCGTAA G C T C C C T A G C	1280
1281	TTCTGGGTTCTTAGTTATTAGAACGAAGATCTAAC C T C C C C G T C C C	1320
1321	AGACCGTTACACTATAACCAAATAAGAAATATAGAAAGTC C T A C T T C G T G C C G T C	1360
1361	CTTCGGAACACCTGGTGGAGCACGGGCCTATTTGGTATC A C T T A A T A A T C C C G	1400

FIGURE 13B

1401	TGTGCATAACAGAAAAATAATATCTATGCCGCTAATGAA	1440
	C G G C C C T C C G	
1441	AATGGTACTATGATCCATTGGCGCCAGAAGATTATACAG	1480
	C C T CCT A C T	
1481	GATTTACTATATGCCAATACATGCCACTCAAGTGAATAA	1520
	C C C T C T C C	
1521	TCAAACCTCGAACATTATTCTGAAAAATTGAAATCAA	1560
	G A C C C C C G C	
1561	GGTGATTCTTAAGATTGAAACAAAGCAACACGACAGCTC	1600
	C G G C G TC T C A	
1601	GTTATACGCTTAGAGGGAATGGAAATAGTTACAATCTTA	1640
	G C TT G C C C C	
1641	TTTAAGAGTATCTCAATAGGAAATTCAACTATTGAGTT	1680
	C G TAGC C T T C C C C T	
1681	ACTATAAACGGTAGAGTTATACTGTTCAAATGTTAATA	1720
	C C ACT T C A C T G C	
1721	CCACTACAAATAACGATGGAGTTAATGATAATGGAGCTCG	1760
	T A G C T C C C C CA	
1761	TTTTCAAGATATTAATATCGGTAATATAGTAGCAAGTGAT	1800
	A CAGC C C C T C C C G CTC C	
1801	AATACTAATGTAACGCTAGATATAAATGTGACATTAAACT	1840
	C C T TT G C C CC C T	
1841	CCGGTACTCCATTGATCTCATGAATATTATGTTGTGCC	1880
	T A C C C	
1881	AACTAATCTTCCACCACTTAT	1902
	C C T T G C	

FIGURE 13C

1	ATGGAGGAAAATAATCAAAATCAATGCATACTTACAATT G C C C T A C	40
41	GTTTAAGTAATCCTGAAGAAGTACTTTGGATGGAGAACG C G C A G T G C T	80
81	GATATCAACTGGTAATTCATCAATTGATATTTCTGTCA C T C C T C C C C C T C	120
121	CTTGTTCAGTTCTGGTATCTAACCTTGACCGGGGGAG T G C C A G C G T T	160
161	GATTTTAGTTGGATTAATAGATTTGTATGGGAATAGT G C C T C C T C C C T C	200
201	TGGCCCTCTCAATGGGATGCATTTCTAGTACAAATTGAA T A C G G G	240
241	CAATTAATTAAATGAAAGAATAGCTGAATTGCTAGGAATG G G C C G G C G C C C C	280
281	CTGCTATTGCTAATTAGAAGGATTAGGAAACAATTCAA C C C G G C T C	320
321	TATATATGTGGAAGCATTAAAGAATGGGAAGAAGATCCT C C G C C G G C	360
361	AATAATCCAGAAACCAGGACCAGAGTAATTGATCGCTTTC C G C C T G G C C A A C A	400
401	GTATACTTGATGGGCTACTTGAAAGGGACATTCCCTCGTT A C T G C C C T G G A T C A C	440
441	TCGAATTTCTGGATTGAAAGTACCCCTTTATCCGTTTAT C A C C C T T C G G C	480
481	GCTCAAGCGGCCAATCTGCATCTAGCTATATTAAAGAGATT A T T C C C C T C A	520
521	CTGTAATTTGGAGAAAGATGGGATTGACAACGATAAA G C C G G C T C	560
561	TGTCAATGAAAACCTATAATAGACTAATTAGGCATATTGAT C G T C C T C C C	600
601	GAATATGCTGATCACTGTGCAAATACGTATAATCGGGGAT G C C C T C C C C T C	640
641	TAAATAATTACCGAAATCTACGTATCAAGATTGGATAAC G C C C T G T T	680
681	ATATAATCGATTACCGAGAGACTAACATTGACTGTATTA C C C A G G A G C C C A T G	720

FIGURE 14A

721	GATATCGCCGTTCTTCCAAACTATGACAATAGGAGAT	760
	C T A C G C	
761	ATCCAATTCAGCCAGTTGGTCAACTAACAAAGGAAGTTA	800
	C T C A G T C A C	
801	TACGGACCCATTAATTAATTAAATCCACAGTTACAGTCT	840
	T C T C C C T G AAG	
841	GTAGCTCAATTACCTACTTTAACGTTATGGAGAGCAGCC	880
	C C C T C A C C TC	
881	GAATTAGAAATCCTCATTATTTGATATATTGAATAATCT	920
	T C G C A C G C C C C	
921	TACAATCTTACGGATTGGTTAGTGGACGCAATTTC	960
	T C C C C G T C C	
961	TATTGGGGAGGACATCGAGTAATATCTAGCCTATAGGAG	1000
	T C A G C C C T C T T	
1001	GTGGTAACATAACATCTCCTATATATGGAAAGAGAGGC	1040
	G T C C T A	
1041	CCAGGAGCCTCCAAGATCCTTACTTTAACGCGTA	1080
	A C TAGT G C C C T A C	
1081	TTTAGGACTTATCAAATCCTACTTTACGATTATTACAGC	1120
	C A C G T C C G A G C C	
1121	AACCTTGGCCAGGCCACCATTAAATTACGTGGTGTGA	1160
	T T C CC TA A	
1161	AGGAGTAGAATTTCTACACCTACAAATAGCTTACGTAT	1200
	G C T G C T C C T C T C	
1201	CGAGGAAGAGGTACGGTTGATTCTTAACTGAATTACCGC	1240
	A T A C C G C C C A	
1241	CTGAGGATAATAGTGTGCCACCTCGCGAAGGATATAGTCA	1280
	A C C CA G C CTCC	
1281	TCGTTTATGTCATGCAACTTTGTTCAAAGATCTGGAACA	1320
	CA G G C C C C G G C T C T	
1321	CCTTTTTAACAACTGGTGTAGTATTTCTGGACCGATC	1360
	A CC C T A A T G C A T	
1361	GTAGTGCAACTCTTACAAATACAATTGATCCAGAGAGAAT	1400
	T C T C C G	

FIGURE 14B

1401	TAATCAAATACCTTAGTGAAGGATTAGAGTTGGGG C C A G C G T C C T G A	1440
1441	GGCACCTCTGTCAATTACAGGACCAGGATTACAGGAGGG A T C C C T	1480
1481	ATATCCTCGAAGAAATACCTTGTTGATTTGTATCTCT T A C T C C G A G C	1520
1521	ACAAGTCATAATTAAATTACCAATTACCCAAAGATAACCGT C T C C C T T T	1560
1561	TTAAGATTCGTTACGCTTCCAGTAGGGATGCACGAGTTA C C G A TTCCC T C TA C	1600
1601	TAGTATTAACAGGAGCGGCATCCACAGGAGTGGGAGGCCA C G C C C A T T C T C T A	1640
1641	AGTTAGTGTAAATATGCCTCTTCAGAAAATATGGAAATA CTCC G C A C G G C	1680
1681	GGGGAGAACCTAACATCTAGAACATTAGATATAACCGATT C G C G C C C C	1720
1721	TTAGTAATCCTTTTCATTTAGAGCTAATCCAGATATAAT CTC C CAGT CC T C C T C C	1760
1761	TGGGATAAGTGAACAACCTCTATTGGTGCAGGTTCTATT C T C C A T AGC C	1800
1801	AGTAGCGGTGAACTTATATAGATAAAATTGAAATTATTC TCATCT C T G C T C G G C	1840
1841	TAGCAGATGCAACATTGAAAGCAGAACATCTGATTTAGAAAG T C C T CC C G T G A C A C C T G	1880
1881	AGCACAAAAGCGGTGAATGCCCTGTTACTTCTCCAAT C G T C C C C CA	1920
1921	CAAATCGGGTTAAAAACCGATGTGACGGATTATCATATTG G C T C G T A C T T C C	1960
1961	ATCAAGTATCCAATTAGTGGATTGTTATCAGATGAATT C G C G C ACC ACC TAGC G	2000
2001	TTGTCTGGATGAAAAGCGAGAATTGTCCGAGAAAGTCAAA C C C C G T C C T	2040
2041	CATGCGAAGCGACTCAGTGTGAGCGGAATTACTTCAAG C C T C C A C C T G	2080
2081	ATCCAAACTTCAGAGGGATCAATAGACAAACCAGACCGTGG CT C A AC C G G A	2120

FIGURE 14C

2121	CTGGAGAGGAAGTACAGATATTACCATCCAAGGAGGAGAT	2160
	T G T C C G G C C C	
2161	GACGTATTCAAAGAGAATTACGTACACTACCGGGTACCG	2200
	T G G C C C T C A T T	
2201	TTGATGAGTGCTATCCAACGTATTTATATCAGAAAATAGA	2240
	C C C T C C G C G C	
2241	TGAGTCGAAATTAAAAGCTTATACCCGTTATGAATTAAGA	2280
	C C C C T C A G C C T	
2281	GGGTATATCGAAGATAAGTCAAGACTTAGAAATCTATTTGA	2320
	C C C C T C C	
2321	TCCGTTACAATGCAAAACACGAAATAGTAAATGTGCCAGG	2360
	A G C G G C C G C	
2361	CACGGGTTCTTATGGCCGCTTCAGCCCCAAATGCCAATC	2400
	T T C C A T T T C T C T	
2401	GGAAAGTGTGGAGAACCGAATCGATGCGCGCCACACCTTG	2440
	G G T C A T T	
2441	AATGGAATCCTGATCTAGATTGTTCTGCAGAGACGGGA	2480
	G C T G C C G T C	
2481	AAAATGTGCACATCATTCCCATCATTCACCTTGGATATT	2520
	G G C C T C T C C	
2521	GATGTTGGATGTACAGACTAAATGAGGACTTAGGTGTAT	2560
	G T C G C C A C	
2561	GGGTGATATTCAAGATTAAGACGCAAGATGGCCATGCAAG	2600
	C C C C C A C	
2601	ACTAGGGAATCTAGAGTTCTCGAAGAGAAACCATTATTA	2640
	T C C T G G C	
2641	GGGGAAGCACTAGCTCGTGTAAAAGAGCGGAGAAGAAGT	2680
	T T C G A A	
2681	GGAGAGACAAACGAGAGAAACTGCAGTTGGAAACAAATAT	2720
	G T C G A G T C	
2721	TGTTTATAAGAGGCAAAAGAATCTGTAGATGCTTTATTT	2760
	C C G C G G C G C	
2761	GTAAACTCTCAATATGATAGATTACAAGTGGATACGAACA	2800
	G C C A G G C C C	
2801	TCGCCATGATTGCGGCAGATAAACCGCGTTCATAGAAT	2840
	C C C C T G C C C	

FIGURE 14D

2841	CCGGGAAGCGTATCTGCCAGAGTTGTCTGTGATTCCAGGT	2880
	T T G T CT T C C T	
2881	GTCAATGCGGCCATTTCGAAGAATTAGAGGGACGTATTT	2920
	G C T C G C T C	
2921	TTACAGCGTATTCTTATATGATGCGAGAAATGTCATTAA	2960
	C A TC G C C C C	
2961	AAATGGCGATTCATAATGGCTTATTATGCTGGAACGTG	3000
	G C T C C CAGC T	
3001	AAAGGTCATGTAGATGTTAGAAGAGCAAAACAACCACCGTT	3040
	G C G G A G T G	
3041	CGGTCCCTGTTATCCCAGAAATGGGAGGCAGAAGTGTACAA	3080
	C G G T G A T C	
3081	AGAGGTTCGTGTCTGTCCAGGTCTGGCTATATCCTTCGT	3120
	A A C T C	
3121	GTCACAGCATATAAGAGGGATATGGAGAGGGCTGCGTAA	3160
	G C T C G C T T G	
3161	CGATCCATGAGATCGAAGACAATAACAGACGAACGTAAATT	3200
	C C GA C C G T G	
3201	CAGCAACTGTGTAGAAGAGGAAGTATATCCAAACAACACA	3240
	TC C C G A A C C C	
3241	GTAACGTGTAATAATTATACTGGGACTCAAGAAGAATATG	3280
	T T C CG C C T A G G C	
3281	AGGGTACGTACACTTCTCGTAATCAAGGGATATGACGAAGC	3320
	GA G C AGC CAG T CA	
3321	CTATGGTAATAACCCCTTCCGTACCAAGCTGATTACGCTTCA	3360
	TCC TCXXXXXXXXXX T T C T C C	
3361	GTCTATGAAGAAAAATCGTATACAGATGGACGAAGAGAGA	3400
	G C G G C C CA C T	
3401	ATCCTTGTGAATCTAACAGAGGGCTATGGGATTACACACC	3440
	C C G TC T CA C	
3441	ACTACCGGCTGGTTATGTAACAAAGGATTAGAGTACTTC	3480
	T A T C T C G C T T	
3481	CCAGAGACCGATAAGGTATGGATTGAGATCGGAGAAACAG	3520
	T C A G C T C	
3521	AAGGAACATTCTCGTGGATAGCGTGGATTACTCCTTAT	3560
	G C C GC T T G	
3561	GGAGGAA	3567

FIGURE 14E

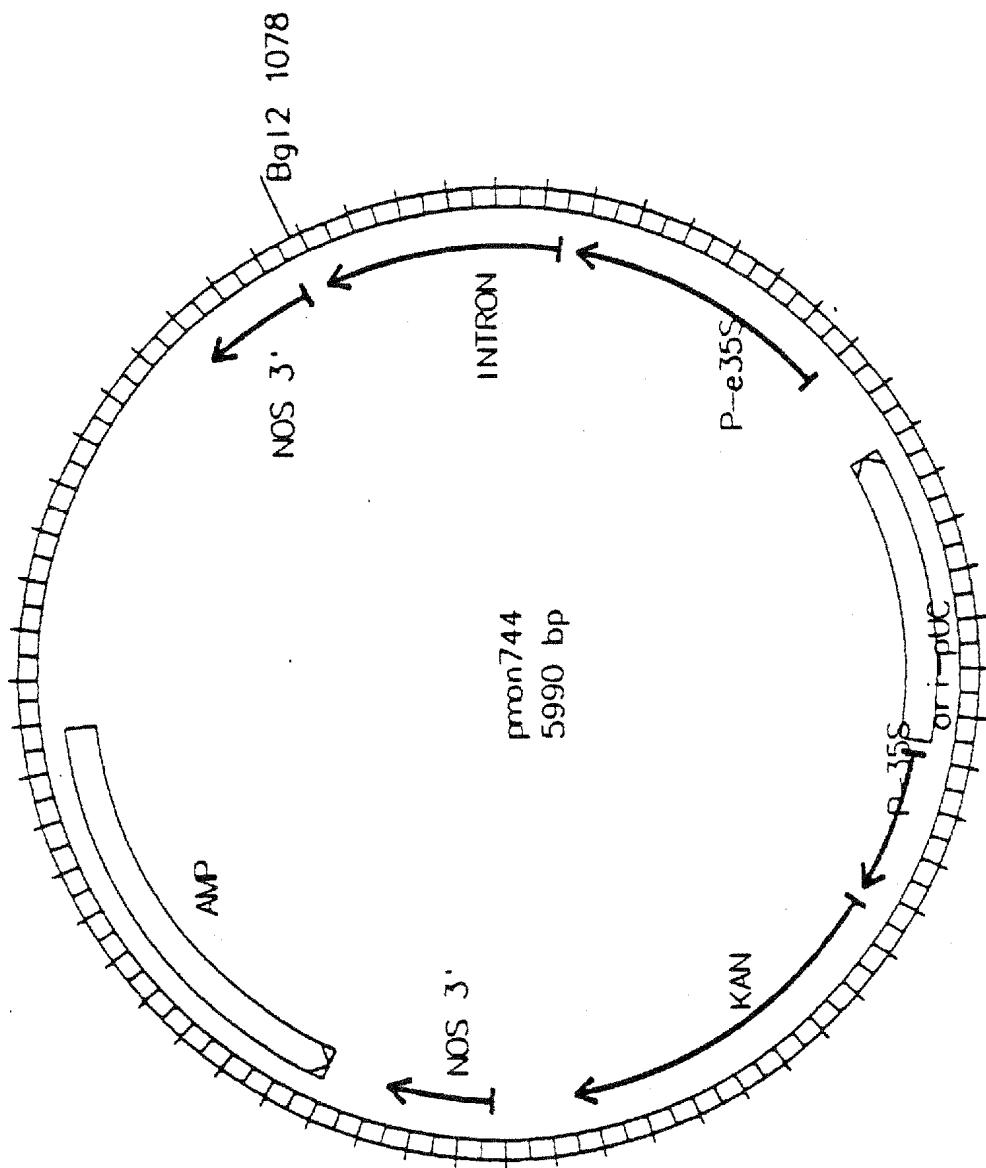


FIGURE 15

1	AGATCTAGAGGTAAATTGTTATGAGTACTGTCGTGGTTAAG GATC	40
41	GGAAACGTCAACGGTGGTGTACAACAAACCTAGAAGGAGGA G T A	80
81	GAAGGCATCCCTCGCAGGAGGGCTAACAGAGTACAGCC T A T	120
121	AGTGGTTATGGTCACTGCTCCTGGCGAACCCAGGAGGAGG GC A A A	160
161	AGACGCAGAAGAGGAGGCAATCGCAGGTCAAGAAGAACTG A G T A	200
201	GAGTTCCCAGGGGAAGGGGCTCAAGCGAGACATTGTGTT A A T	240
241	TACAAAGGACAACCTCGTGGCAACTCCCAAGGAAGTTTC	280
281	ACCTTCGGACCAAGTGTATCAGACTGTCCAGCATTCAAGG T	320
321	ATGGAATACTCAAGGCCTACCATGAGTACAAGATCACAAAG T	360
361	TATCCTTCTTCAGTCGTCAGCGAGGCCTCTTCCACCTCA T G T	400
401	CCAGGATCCATCGCTTATGAGTTGGACCCACATTGCAAAG C A T	440
441	TATCATCCCTCCAGTCCTACGTCAACAAAGTCCAATCAC T	480
481	AAAGGGAGGAGCTAAGACCTATCAAGCTAGGATGATCAAC T T C T	520
521	GGAGTAGAATGGCACGATTATCTGAGGATCAGTGCAGGA T T A	560
561	TACTTGGAAAGGAAGTGGAAAATCTCAGACCCAGCAGG C A G T T	600
601	ATCTTCAGAGTCACCATCAGAGTGGCTCTCAAAACCC T T A	640
641	AAGTAATAGACTCCGGATCAGAGCCTGGTCCAAGCCCACA A T	680

FIGURE 16A

681 ACCAACACCCACTCCA ACTCCCCAAAAGCATGAGCGATTT 720
721 ATTGCTTACGTCGGCATA CCTATGCTGACCATTCAAGAAT 760
761 TC 762

FIGURE 16B